PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES

Volume 13 May 15, 1927 Number 5

A QUANTITATIVE TREATMENT OF DEVIATIONS FROM RAOULT'S LAW

By Joel H. Hildebrand

Chemical Laboratory, University of California

Communicated April 6, 1927

The relationship between the composition of a solution and its various colligative properties has long constituted a physicochemical problem of major importance. Van't Hoff's law for osmotic pressure and Henry's law relating composition to gas pressure or to fugacity are now recognized as failing in concentrated solutions,[1] so that Raoult's law, holding for many solutions throughout the entire range of composition, has been rather generally accepted as defining the ideal solution. It has been possible to predict with substantial accuracy what solutions will obey Raoult's law, and the writer[2] has presented extensive data showing that deviations from this law at any one temperature are approximately proportional to differences in internal pressure. In applying this rule to the solubility of solids in liquids it has not heretofore been possible to make very satisfactory comparisons between the solubilities of different solids because the deviation from the "ideal solubility," based upon Raoult's law, depends upon the composition, the temperature and the melting point of the solute; hence different temperatures for comparison would yield very different values for the deviation.

An approximate quantitative treatment of solubility curves on the basis of their relative deviations from Raoult's law has been given by Mortimer,[3] who considers them as straight lines converging to the melting point, whose slopes represent the heats of solution, which may be calculated from a table of relative internal pressures. While this treatment is practically useful in systems which deviate only moderately from Raoult's law, it suffers from the facts, first, that the curves are not linear, second, that where the deviation from Raoult's law is great enough to yield two liquid phases the curve does not converge with the Raoult's law curve to the melting point, and third, that the slope of the solubility curve has no simple physical meaning, for it corresponds to the heat of solution only where Raoult's law is obeyed.

An examination of the solubility curves for iodine,[3] sulphur,[4] and stannic iodide[5] obtained by plotting log N_2 against $1/T$, where N_2 is the mole fraction of the solute, published by the writer and co-workers, reveals that most of the solvents represented give with each solute a characteristic family of curves, evidently due to simpler relationships than can exist in such divergent cases as iodine in glycerine, or sulphur in liquid ammonia. It is reasonable, therefore, to regard the curves belonging to each family as in a sense regular or normal, and the divergent curves as abnormal. Although there is a great difference between the ideal and actual solubility in the case of iodine in hexane, this fact is not nearly so significant as the evident deviation from the family group in such cases as iodine in glycerine, in benzene, and at low temperatures in carbon disulphide. Deviation from Raoult's law in itself does not interfere with the violet color which these solutions show in common with iodine vapor, while deviation from the family regularity introduces a brown coloration indicating a more specifically chemical effect of solvation.[6] An expression, therefore, which would account for the curves that fall within the family would be of wider significance and applicability than Raoult's law, and would serve to bring into the realm of concentrated solutions all of the physico-chemical laws involving the several colligative properties of such solutions. Deviations from such a law could logically be ascribed to specific chemical effects.

It seems desirable to retain the designation ideal, or perfect, for solutions obeying Raoult's law where, as in the case of the laws for the ideal or perfect gas, one may neglect the effect of specific attractive forces and molecular dimensions, hence we suggest the term *regular* for the solutions which conform to the family of solubility curves, all others being considered *irregular*.

A study of a large proportion of the existing data has revealed the fact that they can be accounted for with surprising accuracy by a rather simple extension of Raoult's law. This law may be written as $f = f^0N$, where f^0 is the fugacity (approximately given by the vapor pressure) of the pure component of a solution, and f is its fugacity in a solution where its mole fraction is N. We may define the activity, a, by the relation $a = f/f^0$, and express Raoult's law as $a = N$ for each component of the solution. Margules suggested a number of years ago that the behavior of a system might be expressed, in general, by writing,

$$a_1 = N_1 e^{\alpha_2 N_2 + \frac{1}{2}\beta_2 N_2^2 + \frac{1}{3}\gamma_2 N_2^3 + \ldots},$$

and

$$a_2 = N_2 e^{\alpha_1 N_1 + \frac{1}{2}\beta_1 N_1^2 + \frac{1}{3}\gamma_1 N_1^3 + \ldots},$$

where the subscripts are added to designate the two components of a mixture. We will confine our discussion to binary mixtures. The ap-

plication of the Duhem equation shows that $\alpha_1 = \alpha_2 = 0$, and also gives the relations between β_1, γ_1, etc., and β_2, γ_2, etc. Where the system can be accounted for by the N^2 terms alone it can be shown that $\beta_1 = \beta_2$. Porter[7] has called attention to the fact that this suffices to describe several systems, and the writer has confirmed it for many others. This corresponds to symmetrical relations between a and N for the two components.[8]

The variation of the exponent of e with temperature is even more important in the solution of our problem than its variation with N. It has been found that the data correspond in most cases to the simple expression,

$$a_2 = N_2 e^{\frac{bN_1^2}{RT}}, \qquad (1)$$

or

$$\ln a_2 = \ln N_2 + \frac{bN_1^2}{RT},$$

while in those cases where the system is unsymmetrical with respect to the two components we may write the more general expression

$$a_2 = N_2 e^{\frac{1}{RT}(bN_1^2 + cN_1^3 + \ldots)}. \qquad (2)$$

It should be noted that the exponential term is equal to the logarithm of the activity coefficient, $\ln \gamma_2 = \ln \frac{a_2}{N_2} = \frac{bN_1^2}{RT}$.

Equation (1) can be tested by noting that the activity of the solute is the same in all saturated solutions at the same temperature, including the solution obeying Raoult's law, where the solubility, which we will write N_2^R, can be calculated from the melting point and the heat of fusion of the solute. We can, therefore, write

$$\ln N_2^R = \ln N_2 + \frac{bN_1^2}{RT}, \qquad (3)$$

and test it by determining the constancy of $k = b/2.303R$. In this brief preliminary account we will give only a few representative calculations.

TABLE 1

DATA FOR SULPHUR

Solvent								
Heptane	t	0	25	35	45	54		
	100 N_2	0.0484	0.1413	0.2005	0.274	0.363		
	k	697	688	689	693	694		
Carbon tetrachloride	100 N_2	0.203	0.500	0.697	0.944	1.212		
	k	529	529	528	527	528		
	t	8	21	39	54	65	72	100
Benzene	100 N_2	0.368	0.557	1.029	1.55	2.18	2.79	6.07
	k	493	498	495	499	497	487	485

Table 1 gives values of k calculated for sulphur in several solvents. It will be seen that in spite of the considerable temperature range the constancy of k is altogether satisfactory. The same result is obtained consistently for all the other regular solubility curves for these substances. The evidence is sufficiently extensive as to justify the wide use of equation (1) to describe what we have called regular solutions. The constant b can be determined from a single point, or can be calculated with fair approximation from internal pressures or from other solubility data involving the same components.

There are various interesting corollaries of equation (1), which may, where necessary, be appropriately extended to the more general equation (2).

1. If we write equation (3) as

$$RT \ln \frac{N_2^R}{N_2} = bN_1^2,$$

we may see that the left-hand member represents the change in the partial molal free energy of component 2 in going from a solution in which it obeys Raoult's law to any regular solution, and that this is not a function of temperature, hence the process is not accompanied by any change in entropy. As a further corollary of this the temperature coefficient of the e.m.f. of a concentration cell having the pure substance at one electrode is given by the simple expression

$$\frac{d\mathbf{E}}{dT} = \frac{R}{\mathbf{NF}} \ln N_2, \qquad (4)$$

even in cases where equation (2) must be used. The data obtained by Taylor[9] for alloy concentration cells over wide ranges of temperature and composition conform very closely to this equation for most of the systems investigated.

2. Differentiating equation (1) with respect to T at constant N gives

$$\left(\frac{\partial \ln a_2}{\partial T}\right)_{N_1} = \frac{-bN_1^2}{RT^2},$$

but the left-hand member is equal also to $\Delta\overline{H}_2/RT^2$, where $\Delta\overline{H}_2$ is the partial molal heat of mixing of component 2, therefore,

$$\Delta\overline{H}_2 = -bN_1^2. \qquad (5)$$

3. Differentiating equation (1) with respect to the pressure upon the liquid phase, P, at constant N and T gives

$$\left(\frac{\partial \ln a_2}{\partial P}\right)_{N,T} = \frac{N_1^2}{RT}\left(\frac{\partial b}{\partial P}\right)_{T,N},$$

but this is also equal to $(\overline{V}_2 - V_2)/RT$, where V_2 is the molal volume of liquid 2 and $\overline{V}_2$ its partial molal volume in the solution; hence

$$\overline{V}_2 - V_2 = N_1^2 \left(\frac{\partial b}{\partial P}\right)_{T,N}. \tag{6}$$

This gives a justification for the equation discovered by the writer some years ago[10] relating the deviation from Raoult's law to $\overline{V}_2 - V_2$. It furthermore harmonizes with a relation discovered by Biron[11] that the expansion of two liquids upon mixing

$$\Delta V = (\overline{V}_1 - V_1)N_1 + (\overline{V}_2 - V_2)N_2 = KN_1N_2,$$

where K is our $(\partial b/\partial P)_{T,N}$.

4. Again we may differentiate equation (1) to get

$$\left(\frac{\partial \ln a_2}{\partial \ln N_2}\right)_{T_1} = 1 - \frac{2bN_1N_2}{RT}. \tag{7}$$

The left-hand member appears in the rigid derivations a number of thermodynamic formulas relating to solutions, and is equal also to the "constant" of Henry's law. This simple formulation of it is of great aid in drawing more accurate conclusions.[12] For example, the heat of solution for any substance forming a regular solution is

$$\Delta H_2 = RT^2 \left(1 - \frac{2bN_1N_2}{RT}\right) \frac{d \ln N_2}{dT}. \tag{8}$$

5. The critical mixing temperature of two liquids, T_c, is that temperature at which[13] $\dfrac{\partial a}{\partial N} = 0$, or $\dfrac{\partial \ln a}{\partial \ln N} = 0$. Applying this criterion to equation (7) we get $RT_c = 2bN_1N_2$. For symmetrical systems the composition of the system at the critical temperature is $N_1 = N_2 = 0.5$, whence $N_1N_2 = 0.25$. Even where the system is not quite symmetrical N_1N_2 at T_c varies but little from 0.25. We may, therefore, write $2RT_c = b$, (9) which allows either T_c or b to be calculated for the other, or the solubility of the solid to be calculated from its critical mixing temperature in the liquid form.

It is interesting to note, in conclusion, that whereas equation (1) occurred to the writer as a result of attempting to fit the facts by giving b the dimensions of free energy, it may be derived also from equations given in 1910 by van Laar[14] on the basis of the van der Waal's equation for gases. If the van der Waal's constants were actually constant, and if certain relationships hold between the components of the mixture, our equation (1) would follow.

[1] Lewis, G. N., *J. Amer. Chem. Soc.*, **30**, 668 (1908).

[2] Hildebrand, J. H., *Solubility*, Chem. Cat. Co., New York, **1924**. Extensive literature references are here given.

[3] Mortimer, *J. Amer. Chem. Soc.*, **44**, 1416 (1922); **45**, 633 (1923).

[4] Hildebrand, J. H., and Jenks, C. A., *Ibid.*, **42**, 2180 (1920); **43**, 2172 (1921); cf. also Ref. 2, p. 148.

[5] Dorfman, M., and Hildebrand, J. H., *Ibid.*, **49**, 729 (1927).

[6] Cf. Hildebrand, J. H., and Glasscock, B. L., *Ibid.*, **31**, 26 (1909).

[7] Porter, A. W., *Trans. Faraday Soc.*, **16**, 336 (1921).

[8] Cf. Ref. 2, p. 48.

[9] Taylor, N. W., *J. Amer. Chem. Soc.*, **45**, 2865 (1923).

[10] Hildebrand, J. H., and Eastman, E. D., *Ibid.*, **37**, 2459 (1915); cf. also Ref. 2, pp. 64–65.

[11] Biron, E., *J. Russ. Phys. Chem. Soc.*, **44**, 1264 (1912).

[12] See, for example, Ref. 2, pp. 146, 161, 194.

[13] Cf. Ref. 2, p. 55.

[14] van Laar, J. J., *Z. physik. chem.*, **72**, 723 (1910).

THE THERMAL DECOMPOSITION OF OZONE

BY OLIVER R. WULF* AND RICHARD C. TOLMAN

GATES CHEMICAL LABORATORY, CALIFORNIA INSTITUTE OF TECHNOLOGY

Communicated April 8, 1927

The kinetics of ozone decomposition possess considerable importance and interest, and much experimental work has been done both on the thermal and the photochemical change. The thermal decomposition has, however, proved bothersome to control experimentally. The authors have been engaged for some time in an investigation of this reaction, and the results of the work are embodied in three articles soon to appear elsewhere. The purpose of this paper is to give a résumé of the work, stating the results and mentioning briefly the most important characteristics of this reaction.

Remarkable discrepancies existed in the results of the seven or more earlier researches. A substantial part of the present task was, therefore, to survey critically this work, and the conclusions finally arrived at come from a consideration of both our own experimental results and the results of earlier investigators.

The thermal decomposition can be made to take place homogeneously at 100° and above in pyrex and in certain other glasses. This appears to be established by the work of Warburg[1] and in particular of Clarke and Chapman[2] as well as from our own results in reaction vessels of different surface to volume ratio. Under certain conditions, reaction at the walls of the container may take place to a considerable extent, as seems probable from the work of Perman and Greaves[3] and of Griffith and McKeown.[4]

The homogeneous decomposition can be carried out so as to be very

closely of the second order with respect to ozone. The earlier careful researches of Warburg and of Jahn[5] showed this. The work of Clarke and Chapman and of Chapman and Jones[6] is in agreement with it. In contrast, however, are the results of Perman and Greaves and also of Griffith and McKeown, who found that the calculated second-order coefficient increased as the reaction progressed. In neither of these cases does it seem certain, however, that the cause was reaction at the walls. In our own work, in eleven experiments made under the most favorable conditions of gas purity, the second-order character is closely adhered to. In other of our experiments, under less favorable conditions, cases were found in which the calculated second-order coefficient varied markedly as the reaction progressed.

The specific rate for different samples of the gas, all of which decompose in accordance with the second order may, nevertheless, vary quite markedly. Thus, in the results of six of our experiments made with the same kind of ozone, the mean deviation from the mean of the specific rate was about 12% of the mean value. Similarly in the results of eight runs recorded by Warburg deviations of approximately the same order were observed. These deviations are apparently greater than could be caused by the errors in the experimental measurements. It appears, in general, that ozone decomposition is very susceptible to the accelerative influence of small amounts of catalytic impurities, although if the cause of the deviations here mentioned is the presence of catalytic impurities, it is rather remarkable that the order remains so nearly the second over the larger portion of the complete decomposition. From the many results in which the reaction has been observed to proceed according to the second order an idea of the numerical value of the specific rate may be obtained. If there is a true second-order rate for pure ozonized oxygen, it has at 100°C. and 1 atmosphere total pressure a specific rate as low as $1.5\text{–}1.7 \times 10^2$ cc./-(mol. sec.).

During the decomposition of the ozone in the common ozonized oxygen of low per cent ozone the total pressure increases only slightly, and is for most purposes practically constant. When, however, the total pressure is materially altered, in other ways, the specific second-order rate is found to change. For ozonized oxygen, which decomposes in accordance with the second order, the specific rate is closely inversely proportional to the total pressure. This was originally demonstrated by Jahn and this influence of total pressure has been observed by others including ourselves.

For this rather surprising fact two possible explanations present themselves. (1) The decomposition of ozone is inhibited by oxygen, the specific second-order rate being inversely proportional to the concentration of oxygen. This hypothesis was believed by Jahn to be true, and was the fact used by him to support the particular mechanism which he proposed

for ozone decomposition. This mechanism attributed ozone decomposition to the reaction of molecular ozone with atomic oxygen, the latter, with molecular oxygen, being in thermodynamic equilibrium with molecular ozone. (2) The decomposition of ozone is inhibited by some unknown negative catalyst present in ozonized oxygen and the specific second-order rate is inversely proportional to the concentration of this negative catalyst, so that the specific rate is inversely proportional to total pressure.

Experiments on the effect of total pressure are clearly not sufficient to differentiate between these two hypotheses. Jahn and Chapman and Jones had each made some measurements that were sufficient in principle to distinguish. The results, however, were contradictory; those of Jahn indicating an inhibiting effect due to oxygen closely proportional to its concentration, while those of Chapman and Jones showed oxygen to possess at the most only a small inhibiting effect. The question as to which of these two hypotheses was in reality correct was thus left unanswered. During the present investigations experiments have been made attempting particularly to answer this question, and the results have led to the conclusion that the decrease in the specific second-order decomposition rate which accompanies increase in the total pressure of oxygen-ozone mixtures is due to the inhibiting effect of oxygen on the decomposition.

In general, oxygen is liable to contain positive catalyst for the decomposition. Different samples of oxygen show different degrees of ozonizability in an electric discharge ozonizer, and there is a tendency for high ozonizability to be correlated with low specific second-order decomposition rate. By electrolysis under suitable conditions from sulphuric acid, oxygen and ozone can be prepared which contain either no catalyst for ozone decomposition or at least a nearly constant amount.

Finally, in the present investigations, the temperature coefficient[7] of the rate of this decomposition has been measured at one atmosphere total pressure, using five temperatures in the range 148–179°C. The rate of decomposition may be expressed as a function of the temperature at *one atmosphere* pressure by the equation

$$-\frac{dC}{dt} = 2.04 \times 10^{20}\, e^{-\frac{30,900}{RT}}\, C^2.$$

The rate under the same conditions may also be expressed by the equation

$$-\frac{dC}{dt} = \frac{5.89 \times 10^{13}}{C_{O_2}}\, T^{\frac{1}{2}}\, e^{-\frac{29,600}{RT}}\, C^2$$

and this equation will also give reasonable values at total pressures other than one atmosphere.

These equations are to be regarded as purely empirical expressions. If, however, it be assumed that the decomposition actually occurs in ac-

cordance with the mechanism proposed by Jahn, the temperature coefficient leads to a value for the heat of dissociation of ozone into molecular oxygen and atomic oxygen. Calculating by thermodynamic methods the concentration of monatomic oxygen at 100° and 1 atmosphere pressure in 5% ozone, and by kinetic theory methods the number of collisions between molecular ozone and atomic oxygen, it has been shown that these collisions are many times too small in number to account for the observed rate and hence that the Jahn mechanism cannot be regarded as tenable, at least in its original simple form.

* NATIONAL RESEARCH FELLOW IN CHEMISTRY.

[1] Warburg, *Ann. Physik*, **9**, 1286 (1902).

[2] Clarke and Chapman, *J. Chem. Soc.*, **93**, 1638 (1908).

[3] Perman and Greaves, *Proc. Roy. Soc.*, **A80**, 353 (1908).

[4] Griffith and McKeown, *J. Chem. Soc.*, **127**, 2086 (1925).

[5] Jahn, *Zeit. anorg. Chemie*, **48**, 260 (1905).

[6] Chapman and Jones, *J. Chem. Soc.*, **97**, 2463 (1910).

[7] Earlier observations of the influence of temperature on the specific rate had been made by Warburg (loc. cit.); Clement, *Ann. Physik*, **14**, 341 (1904); and Perman and Greaves (loc. cit.); and recently by Belton, Griffith and McKeown, *J. Chem. Soc.*, **129**, 3153 (1926).

CONCERNING POINTS OF A CONTINUOUS CURVE THAT ARE NOT ACCESSIBLE FROM EACH OTHER

BY C. M. CLEVELAND

DEPARTMENT OF APPLIED MATHEMATICS, UNIVERSITY OF TEXAS

Communicated April 4, 1927

The purpose of this paper is to establish the following theorem:

THEOREM. *If A and B are two distinct points of a continuous curve M and every simple continuous arc from A to B contains at least one point of M distinct from A and from B then there exists a simple closed curve which is a subset of M and which separates A from B.*

Proof. By a theorem due to Schoenflies,[1] A and B do not both belong to the boundary of the same complementary domain of M. Suppose that A does not belong to the boundary of the unbounded complementary domain of M. Let C denote a circle with center at A and not containing or enclosing B. Since M is regular there exists a circle C_1 with center at A and such that any point of M within C_1 can be joined to A by an arc of M lying within C. There are not[1] more than a finite number of complementary domains of M having A on their boundaries and containing points without C_1. If there be any, let $D_1, D_2, D_3, \ldots, D_n$ denote them, let $T_1, T_2, T_3, \ldots, T_n$ denote their respective boundaries and let H denote

the set $T_1 + T_2 + T_3 + \ldots + T_n$. Let K denote the set of all points X of M such that X can be joined to A by a connected subset of M which contains no point outside of C and let E denote the sum of the sets H and K. Let D denote the complementary domain of E which contains B and let L denote the boundary of D. Since K is[2] a continuous curve and each T_i is[3] a continuous curve, therefore, E is a continuous curve. Hence L is[3] a continuous curve.

Now A does not belong to L. For suppose it does. Then since every point of L is[1] accessible from D there exists an arc AB which, except for A, is a subset of D. Let P denote the first point that the arc AB has in common with C_1. The interval AP of AB has only the point A in common with M and hence $AP-A$ lies in one of the complementary domains $D_1, D_2, D_3, \ldots D_n$. This, however, is impossible since the interval BP of AB contains no point of the boundary of any one of these domains. Hence A lies in one complementary domain of the continuous curve L and B lies in another one. Hence by a theorem of R. L. Moore[4] there exists a simple closed curve which separates A from B and which is a subset of L and, therefore, of M.

[1] Schoenflies, A., "Die Entwickelung der Lehre von den Punktmannigfaltigkeiten," *Zweite Teil*, Leipzig, 1908.

[2] Gehman, H. M., *Ann. Math.*, **27**, 1925 (29–46), Lemma B.

[3] Moore, R. L., *Math. Zs.*, **15**, 1922 (259).

[4] Moore, R. L., loc. cit. (260).

RECENT PROGRESS OF INVESTIGATIONS BY SYMBOLICAL METHODS OF THE INVARIANTS OF BI-TERNARY QUANTICS

BY O. E. GLENN

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF PENNSYLVANIA

Communicated March 21, 1927

The celebrated mathematical discipline, the symbolical theory of invariants of algebraical forms, has its roots in the inspired work of Clebsch. Up to recent times contributions by this method were largely a product of research in Germany.

In the problem determined by forms in two contragredient sets of three variables (x), (u), the object of study is the fundamental system, S, of the connex, $f = a_x^m \alpha_u^n$. As in the analogous theory of binary forms three objectives may be sought. The first is the method used for the generation of the concomitants, which must be definitive for the formation of all invariants of the existing infinitude. Second, finite expansions analogous to Gordan's series and an appropriate theory of symbolic moduli should be

discovered for use in reducing all invariants in terms of a finite set. Third, the complete system thus found may be investigated from the geometrical point of view, for every such system is the algebraic prototype of a projective geometry of plane loci.

Formation of Invariant Connexes.—If neither exponent m, n is zero the locus of f is a connex. In a joint paper[1] published in 1869, in connection with some general symbolical theory of connexes, Clebsch and Gordan determined the complete system of eight concomitants of the bilinear case $f = a_x\alpha_u$. The subject then remained stationary until 1916 when the present writer published a definitive method[2] of generation for connexes. This is a process of transvection between two forms, analogous to binary transvection as due to Cayley and Gordan. The algorithm is as follows: With two quantics assigned,

$$\begin{aligned} \varphi &= Aa_{1x}a_{2x} \ldots a_{rx}\alpha_{1u}\alpha_{2u} \ldots \alpha_{su}, \\ \psi &= Bb_{1x}b_{2x} \ldots b_{\rho x}\beta_{1u}\beta_{2u} \ldots \beta_{\sigma u}, \end{aligned} \tag{1}$$

polarize φ by the operator,

$$\Sigma \left(y_1^{(1)}\frac{\partial}{\partial x}\right)^{\iota_1}\left(y_2^{(1)}\frac{\partial}{\partial x}\right)^{\iota_2} \cdots \left(y_\sigma^{(1)}\frac{\partial}{\partial x}\right)^{\iota_\sigma}\left(y_1^{(2)}\frac{\partial}{\partial x}\right)^{\kappa_1}\left(y_2^{(2)}\frac{\partial}{\partial x}\right)^{\kappa_2} \cdots \left(y_\rho^{(2)}\frac{\partial}{\partial x}\right)^{\kappa_\rho}$$
$$\times \left(v_1^{(1)}\frac{\partial}{\partial u}\right)^{\lambda_1}\left(v_2^{(1)}\frac{\partial}{\partial u}\right)^{\lambda_2} \cdots \left(v_\rho^{(1)}\frac{\partial}{\partial u}\right)^{\lambda_\rho}\left(v_1^{(2)}\frac{\partial}{\partial u}\right)^{\mu_1}\left(v_2^{(2)}\frac{\partial}{\partial u}\right)^{\mu_2} \cdots \left(v_\sigma^{(2)}\frac{\partial}{\partial u}\right)^{\mu_\sigma},$$

where,

$$\left(y_c^{(e)}\frac{\partial}{\partial x}\right) = y_{1c}^{(e)}\frac{\partial}{\partial x_1} + y_{2c}^{(e)}\frac{\partial}{\partial x_2} + y_{3c}^{(e)}\frac{\partial}{\partial x_3}.$$

The terms of the summation are uniquely given by the sets of solutions of the restricted system of linear diophantine equations,

$$\sum_{t=1}^{\sigma} \iota_t = i, \qquad \sum_{\tau=1}^{\rho} \kappa_\tau = j, \qquad \sum_{\tau=1}^{\rho} \lambda_\tau = k, \qquad \sum_{t=1}^{\sigma} \mu_t = l.$$

These are subject to the conditions,

$$i + j \leqq r, \qquad k + l \leqq s, \qquad i + l \leqq \sigma, \qquad j + k \leqq \rho,$$

and are to be satisfied only by values 0,1 of the variables, i, j, k, l, being fixed chosen integers. With t,τ assigned at least one number in each pair (ι_t, μ_t), $(\kappa_\tau, \lambda_\tau)$, must be zero. We substitute in the resulting polar as follows:

$$\begin{aligned} y_p^{(1)} &= \beta_p, \ (p = 1, \ldots, \sigma); \qquad y_p^{(2)} = (b_p u), \ (p = 1, \ldots, \rho), \\ v_p^{(1)} &= b_p, \ (p = 1, \ldots, \rho); \qquad v_p^{(2)} = (\beta_p x), \ (p = 1, \ldots, \sigma). \end{aligned}$$

Finally, multiply each term of the result by the b_x, β_u factors (of ψ), which are not then involved in it. The resulting concomitant is the transvectant,

$$T = (\varphi, \psi)^{ij}_{kl}. \qquad (2)$$

If two infinite systems, $[\varphi]$, $[\psi]$, of connexes (1), both have the technical properties of finiteness and completeness, then the system $[C]$ derived by transvection between $[\varphi]$ and $[\psi]$ is finite and complete.

By $[C]$ is meant the system which consists of all terms of all transvectants T, where φ is a form chosen from $[\varphi]$ and ψ is a form chosen from $[\psi]$. The proof of this theorem gave opportunity for the solution of two problems of enumeration, viz., the simultaneous system of a conic $F = p_x^2$ and the connex $f = a_x\alpha_u$, and the simultaneous system of two connexes $f = a_x\alpha_u$, $\varphi = p_x\theta_u$. The system of F and f was given in my paper of 1916 and consists of 69 forms. For illustration we give the simplest category of these, that in which the order plus the class of the product of the two forms in a transvectant is 4 ($f_1 = b_\alpha\beta_u a_x$, $L = (pqu)^2$).

$$(F, f)^{10}_{00} = p_\alpha p_x a_x, \quad (F, f)^{01}_{00} = (pau)\alpha_u p_x, \quad (F, f)^{11}_{00},$$
$$(F, f_1)^{10}_{00}, \quad (F, f_1)^{01}_{00}, \quad (F, f_1)^{11}_{00}, \quad (L, f)^{00}_{10} = (pqa)(pqu)\alpha_u,$$
$$(L, f)^{00}_{01}, \quad (L, f)^{00}_{11}, \quad (L, f_1)^{00}_{10}, \quad (L, f_1)^{00}_{01}, \quad (L, f_1)^{00}_{11}.$$

I published the system of the pair (f, φ) in 1926. It consists of 193 quantics which were expressed as transvectants.[3]

Bi-Ternary Reducing Series.—Several mathematicians have proposed forms of generalization[4] of Gordan's binary series. In a memoir which is soon to be published the present writer has identified a new algorithm in the symbolic theory which gives a reducing series for bi-ternary transvectants, approximating in generality and utility to the binary series[5] of Stroh, for which the claim was made that it would give all syzygetic relations between multiple transvectants on three forms. We shall give an example of this series and the general formula. By a method of contraction which applies to all symbolical monomials we can show that a form,

$$Q = a_\alpha c_\beta (abc)\gamma_u b_x,$$

is a term of three transvectants, in general, of as many transvectants as there are symbolic pairs (a, α), in which the first form is of degree one. Thus Q is a term of

$$T = (g, (f, f)^{01}_{10})^{10}_{00}, \qquad S = (f, (f, g)^{01}_{00})^{10}_{10},$$

where $f = a_x^2\alpha_u$, $g = a_\alpha a_x$. There is, then, a theorem which was suggested by Gordan[6] but not utilized by him in connection with transvectants, that the difference between a transvectant and one of its terms is a linear combination of transvectants of lower grade number (cf. (1)), $N = r +$

$s + \rho + \sigma$, of forms obtained from the original forms (in T) by convolution. Applying this theorem to both T and S we get two values of Q in the form of series, whence,

$$T - \frac{1}{2}T_1 + \frac{1}{4}T_2u_x + \frac{1}{4}T_3 = S - \frac{1}{2}S_1 + \frac{1}{4}S_2 + \frac{1}{4}S_3, \qquad (3)$$

$$T_1 = (g, c_\beta b_x(\overline{bc}\gamma x))^{01}_{00}, \qquad T_2 = (g, c_\beta(\overline{\overline{bbc}}\gamma u))^{10}_{00},$$

$$T_3 = (g, c_\beta(\overline{bc}\gamma bu)^{00}_{00}, \qquad S_1 = (f, a_\alpha c_x(\overline{ca}\gamma x))^{01}_{10},$$

$$S_2 = (f, a_\alpha(\overline{\overline{cca}}\gamma u)^{10}_{00}, \qquad S_3 = (g, a_\alpha(\overline{\overline{ca}}\gamma cu))^{00}_{00}.$$

When these methods are applied in the general case of three arbitrary connexes,

$$f = a^m_x\alpha^n_u, \qquad g = b^u_x\beta^v_u, \qquad h = c^\sigma_x\gamma^\tau_u,$$

with

$$T = (h,(f, g)^{ij}_{kl})^{pq}_{rs}, \qquad \xi = \sigma_1 + s_2 + \tau_2 + t_1,$$

$$S = (f,(-1)^\xi (h, g)^{\sigma_3 s_3 + \sigma_1 i + \sigma_2 + t_1 j + s_2 - \sigma_1}_{t_3 \tau_3 + t_1 k + \sigma_2 + t_1 l + \tau_2 - t_1},$$

$$\sigma_1 + \sigma_2 + \sigma_3 = p, \quad s_2 + s_3 = q, \quad \tau_2 + \tau_3 = s, \quad t_1 + t_2 + t_3 = r,$$

we obtain the series

$$T + \sum I_1 (\overline{\overset{00a}{h}, \overline{(f, g)^{ij}_{kl}}}^{bcd})^{p_1q_1}_{r_1s_1} u^{m_1}_x$$

$$= S + \sum I_2(\overline{\overset{00a}{f}, (-1)^\xi \overline{(h, g)^{\sigma_3 s_3 + \sigma_1}_{t_3\tau_3 + t_1}}}^{\beta\gamma\delta})^{pqq_2}_{rss_2} u^{m_2}_x. \qquad (4)$$

The numbers I_1, I_2 are rational. The form $\overline{\overset{abc}{P}}$ indicates application, in a product P, of a convolutions which replace a product p_xq_x by (pqu), b convolutions which replace $\alpha_u\beta_u$ by $(\alpha\beta x)$ and c convolutions which convert $p_x\theta_u$ into p_θ. The summations are finite and are arranged according to decreasing grades of the transvectants.

This series adds greatly to our knowledge of the generalities of the subject. Its practical significance is that the series itself and rules of procedure for work of reduction, derived from the principles involved, enable us to find, from the set of fundamental concomitants of degrees $\leqq i-1$ of $f = a^m_x\alpha^n_u$, the complete set of degree i, although the work is complicated in the case of a high degree. We give, in the memoir, a complete system of 32 fundamental concomitants of degrees 1, 2, 3, of $f = a^2_x\alpha_u$.

Relatively Complete Systems.—The following theorem was proved and applied in obtaining a system of 90 forms for $f = a^2_x\alpha_u$, relatively finite and complete with respect to a set of symbolical moduli, and also to delimit by theoretical considerations the absolutely finite and complete system. It corresponds to the principal proposition in Gordan's third proof of Gordan's theorem.

LEMMA. *If a finite system* (A) *of connexes* $A_1, A_2, \ldots$, *all of which are concomitants of* $f = a_x^m \alpha_u^n$, *includes* f *and is relatively complete for a set* (G) *of constant symbolic moduli,* $G_1, G_2, \ldots$, *and also if a finite system* (B) *of connexes,* $B_1, B_2, \ldots$, *concomitants of* f, *is relatively complete for a set* (Γ) *of symbolical moduli,* $\Gamma_1, \Gamma_2, \ldots$, *and contains the leading form of the form-sequence whose quantics have as their only constant symbolical factor the modulus* G_i, *for each case* $i = 1, 2, \ldots$, *then the system* (C) *derived by transvection from* (A) *and* (B) *is relatively finite and complete modulo* (Γ).

In reference to geometry a paper[7] by Dr. Sensenig is a branch of this investigation. He used the transvectant (2) in deriving the complete system of the involution of conics $a_x^2 + kb_x^2$ taken simultaneously with the conic, $(\alpha\beta x)^2$, which is harmonic to a_x^2, b_x^2. The system was reduced in terms of the simultaneous system of a_x^2, b_x^2.

[1] *Math. Ann.*, **1** (1869), 359.

[2] *Trans. Amer. Math. Soc.*, **17** (1916), 405.

[3] *Amer. J. Math.*, **48** (1926), 45.

[4] Study, *Meth. zur. Th. der Ternären Formen*, 1889, 54; Capelli, *Giorn. Battaglini*, **18** (1880), 17; E. Noether, *J. Math.*, **134** (1908), 29.

[5] *Math. Ann.*, **31** (1888), 448.

[6] *Ibid.*, **1** (1869), 90.

[7] *Amer. J. Math.*, **41** (1919), 111.

A DYNAMICAL THEORY OF ECONOMIC EQUILIBRIUM

BY C. F. ROOS*

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CHICAGO

Communicated March 17, 1927

1. *Introduction.*—In a previous paper I have shown how the dynamic problem of competition is related to a dynamic generalization of the static theory of economic equilibrium of Walras and Pareto.[1] In this paper I intend to simplify the notation of my previous paper and develop the theory of dynamic economic equilibrium for functional equations of demand and supply.[2] I will show that the problem of equilibrium for a coöperative society is a problem for which it is desired to maximize a functional operator, and, also, that the problem for a competitive society is a problem for which it is desired to obtain partial maxima of several functional operators.[3]

Inasmuch as I have already given the economic setting of the problem, I shall not attempt to repeat it in detail.[4] It is sufficient for the purposes of this paper to say that in any economic system certain goods and services are transformed into other goods and services and that we wish to

discover a mechanism which will describe this situation. Fortunately, we do not have to grope in the dark to discover this mechanism, for Pareto has described it for a fixed time $t = t_0$ and has shown that there are as many equations of demand, supply, production, etc., as there are variables necessary to define the system and, furthermore, that in general these equations are independent.[5]

2. *Notation.*—Let us suppose that the total number of all commodities produced in unit time at any time t is $\gamma + 1$, and, further, that the total number of services of land, of persons and of capital required to produce these commodities is q. Let us denote these services by the symbols $C_1, \ldots, \ldots, C_q$ and the commodities by $C_{q+1}, \ldots, C_{q+\gamma+1}$. We may then represent services and commodities by C_j $(j = 1, \ldots, q + \gamma + 1)$, it being understood that the first q of the C_j represent services and the last $\gamma + 1$ represent commodities.

If we choose C_{m+1} $(m = r + q)$, as a money unit in terms of which the prices of services and commodities may be represented, we may represent the respective prices for the commodities and services as $p_1(t), \ldots, p_m(t)$.

In general, there will be more than one producer of each commodity and service. Let us suppose that the sequence $\xi_1, \ldots, \xi_{m+1}$ be so chosen that ξ_μ $(\mu = 1, \ldots, m)$ represents the total number of producers of C_μ. We may then represent the rates of production by $u_{\mu k}(t)$ where $k = 1, \ldots, \xi_\mu$ and $\mu = 1, \ldots, m$.

3. *Equations of Demand and Supply.*—Since the prices of the C_μ are expressed in terms of C_{m+1}, there will be m equations giving the rate of demand for the m quantities C_μ and, as I have already pointed out, this rate of demand will be a functional of the rates of production and prices.[1] We may suppose, then, that the rate of demand D_μ for C_μ is given by

$$D_\mu = D_\mu(u, p, t) \qquad (\mu = 1, \ldots, m); \qquad (t_0 \leq t \leq t_1) \qquad (1)$$

where the D_μ are continuous functions of t and continuous functionals of the set (u, p) which is used to denote the set $(u_{11}, \ldots, p_m)$ of $n = m(1 + \sum_{\mu=1}^{m} \xi_\mu)$ variables.

As a likely form of D_μ we might write

$$D_\mu = G_\mu(u, u', p, p', t) + \int_a^b \sum_{\mu=1}^{m} \lambda_\mu e^{\kappa_\mu(t-\tau)} p_\mu(\tau) d\tau \quad (t_0 \leq a \leq t \leq b \leq t_1) \qquad (2)$$

where G_μ is continuous in u, u', p, p' and t, and λ_μ and K_μ are constants.

Corresponding to this first group of m equations of demand there will

be m equations of supply which may be represented by the continuous functionals

$$S_\mu = S_\mu(u, p, t) \qquad (\mu = 1, \ldots, m); \qquad (t_0 \leq t \leq t_1). \quad (3)$$

We may suppose that the relation between the demand and supply is of the form

$$\sum_{\nu=1}^{m} C_{\nu\mu} D_\nu - S =_\mu 0 \qquad (\nu = 1, \ldots, m);\ t_0 \leq t \leq t_1 \quad (4)$$

where the $C_{\nu\mu}$ are continuous functions of the time which at any particular time $t = t_0$ reduce to the coefficients of production of the static theory.[5] We may write the equations (4) as

$$\varphi_\mu(u, p, t) = 0 \qquad (\mu = 1, \ldots, m); \qquad t_0 \leq t \leq t_1 \qquad (5)$$

where the φ_μ are continuous functional operators with respect to (u, p) for every t on (t_0, t_1).

4. *Cost of Production and Profit.*—In general, the cost of producing one unit of C_μ per unit time for the kth producer will depend upon the rates of production, the prices and the time derivatives of these quantities. We may then suppose the cost of production to be

$$\theta_{\mu k} = \theta_{\mu k}(u, u', p, p', t), \qquad (t_0 \leq t \leq t_1) \qquad (6)$$

where μ and k have the same meaning as before.[1]

The total profit, $\pi_{\mu k}$, which the kth producer of C_μ receives for the period of time $a \leq t \leq b$ where $t_0 \leq a \leq b \leq t_1$ will then be

$$\pi_{\mu k} = \int_a^b [p_\mu - \theta_{\mu k}(u, u', p, p', \tau)] u_{\mu k}(\tau) d\tau \begin{Bmatrix} k = 1 \ldots, \xi_\mu \\ \mu = 1, \ldots, m \end{Bmatrix}. \quad (7)$$

These profits, $\pi_{\mu k}$, are continuous functionals.

If we knew the profits, $\pi_{\mu k}$, the equations (5) and (7) would be sufficient to determine all rates of production and prices necessary to define the economic system, and the problem would be solved. In the static theory Walras supposes that the price of C_μ is the cost of production of C_μ, or, that is, that there is no profit.

Let us search for a more general hypothesis than that of Walras.[7]

5. *The Maximum of a Functional.*—As a likely hypothesis we might suppose that the rates of production and prices are determined so that the sum of the profits

$$\sum_{\mu=1}^{m} \sum_{k=1}^{\xi_\mu} \pi_{\mu k}$$

will be as large as possible. This is equivalent to supposing that the economic society is a coöperative society.

For simplicity in notation let us use a single subscript $\rho = 1, \ldots, r$ where $r = m\left(\sum_{\mu=1}^{m} \xi_\mu\right)$ to renumber the profit integrals (7). We may suppose the profits to be given by the r continuous functional operators $W_\rho(y_i, a, b)$ where y_i $(i = 1, \ldots, m + r)$, represents the set (u, p).

The problem of coöperation is, then, that of:

Finding a set of $r + m$ functions $y_i(t)$ of class C^1 which satisfy the conditions

$$\begin{aligned} |\, y_i(t) - y_{i0}(t) \,| < h, \qquad & |\, y_i'(t) - y_{i0}'(t) \,| < h \\ |\, a - a_0 \,| < h, \qquad & |\, b - b_0 \,| < h \end{aligned} \tag{8}$$

and for which

$$\varphi_\mu(y_i, a, b, t) = 0 \qquad (\mu = 1, \ldots, m); \qquad (t_0 \leq t \leq t_1) \tag{9}$$

such that the functional representing profit of the society,

$$W(y_i, a, b) = \sum_{\rho=1}^{r} W_\rho(y_i, a, b),$$

is a maximum.

Inasmuch as this problem is similar to a problem discussed by Hahn, we need only give the hypotheses and results of his paper.[8] In addition to the above hypotheses Hahn supposes further that

(*a*) If

$$y_i(t) = y_{i0}(t) + \sum_{\sigma=1}^{s} \epsilon_\sigma \eta_{i\sigma}(t) + (\epsilon),$$

$$y_i'(t) = y_{i0}'(t) + \sum_{\sigma=1}^{s} \epsilon_\sigma \eta_{i\sigma}(t) + (\epsilon)$$

$$a = a_0 + \sum_{\sigma=1}^{s} \epsilon_\sigma \alpha_\sigma + (\epsilon)$$

$$b = b_0 + \sum_{\sigma=1}^{s} \epsilon_\sigma \alpha_\sigma + (\epsilon)$$

where (ϵ) represents a function of the parameters $\epsilon_1, \ldots, \epsilon_s$ which tends to zero with these parameters, and where $\eta_i\sigma$, α_σ and β_σ are so chosen that the conditions (8) are satisfied, then

$$\begin{aligned} W(y_i, a, b) &= W(y_{i0}, a_0, b_0) + \sum_{\sigma=1}^{s} \epsilon_\sigma V(\eta_{i\sigma}, \alpha_\sigma, \beta_\sigma) + (\epsilon), \\ \varphi_\sigma(y_i, a, b, t) &= \varphi\mu(y_{i0}, a_0, b_0, t) + \sum_{\sigma=1}^{s} \epsilon_\sigma \psi_\mu(\eta_{i\sigma}, \alpha_\sigma, \beta_\sigma, t) + (\epsilon) \end{aligned} \tag{10}$$

in which $V(\eta_i, \alpha, \beta)$ and $\psi_\mu(\eta_i, \alpha, \beta, t)$ are continuous linear functional operators with respect to η_i and η_i' and continuous linear functions of α, β and of α, β, t, respectively.[6]

(*b*) If s sets of functions $\eta_{i\sigma}(t)$ of class C'' on (a_0, b_0) and s sets of constants α_σ, β_σ are given which satisfy the linear functional equations $\psi_\mu(\eta_i, \alpha, \beta) = 0$, then there exists for every set of values of the parameters $\epsilon_1, \ldots, \epsilon_s$ sufficiently small in absolute value, a solution y_i, a, b of the functional equations $\varphi_\mu(y_i, a, b, t) = \varphi_\mu(y_{i0}, a_0, b_0, t)$ of the form (10).

(*c*) For every $e > 0$, there exists a d, such that if $\omega_\mu(t)$ are continuous functions of t on (t_0, t_1), for which $|\omega_\mu(t)| < d$, then there exists at least one set of functions $\eta_i(t)$ of class C^1 on (a_0, b_0) and a pair of constants α, β such that

$$|\eta_i| < e, \qquad |\eta_i'| < e, \qquad |\alpha| < e, \qquad |\beta| < e$$

and such that

$$\psi_\mu(\eta, \alpha, \beta, t) = \omega_\mu(t).$$

With these hypotheses Hahn uses Riesz' theorem to prove the following theorem:

THEOREM 1. *If a set of functions y_i furnishes a solution of the problem, then there exists a constant l_1 and m functions of limited variation λ_μ such that*

$$l_1 V(\eta_i, \alpha, \beta) + \int_{t_0}^{t_1} \sum_{\mu=1}^{m} \psi_\mu(\eta_i, \alpha, \beta, t) d\lambda_\mu(t) = 0 \qquad (11)$$

for every set of functions η_i of class C^1 in a_0, b_0 and for all numbers α, β.

It is worth mentioning here that this generalized Euler-Lagrange rule includes as special cases the ordinary Lagrange problems in the calculus of variations.

6. *Partial Maxima of Functionals.*—A type of society which is more common than the coöperative society is the competitive society. In such a society each producer considers the rates of production of his competitors as fixed and determines his own rate of production so that his profit is as large as possible. The problem of competition must, therefore, be considered as a problem for which it is desired to find the partial maxima of several functional operators.[3] We may state the problem mathematically as follows:

Choose the rate of production y_ρ $(1 \leq \rho \leq r)$ of the ρth producer and the prices satisfying the equations

$$\varphi_\mu(y_i, a, b, t) = 0 \qquad (i = 1, \ldots, r, r+1, \ldots, r+m); \ (t_0 \leq t \leq t_1)$$

in which y_κ $(\kappa = 1, \rho - 1, \rho + 1, \ldots r)$ are held constant such that each

$$W_\rho(y_i, a, b)$$

is a maximum when the y_κ are held constant.

Let us suppose that the conditions (*a*), (*b*) and (*c*) are satisfied for the partial variations $\eta_{\rho i}$ of the y_ρ when the y_κ are held constant. The subscript i must then be supposed to take on successively the values $1, r+1, \ldots, r+m$; $2, r+1, \ldots, r+m$; $\ldots$; $r, r+1, \ldots, r+m$ when ρ takes on the successive values $1, 2, \ldots, r$.

The problem of competition can now be solved by an analysis similar to that of Hahn, in fact, we have the following theorem:

THEOREM 2. *If a set of functions* $y_1, \ldots, y_r, y_{r+1}, \ldots, y_{r+m}$ *furnishes a solution of the problem of partial maxima of functionals, then there exists a set of* r *constants* l_ρ *and* r *sets of* m *functions of limited variation* $\lambda_{\rho\mu}$ *such that*

$$l_1 V_1(\eta_{1i}, \alpha, \beta) + \int_{t_0}^{t_1} \sum_{\mu=1}^{m} \psi_\mu(\eta_{1i}, \alpha, \beta, t) d\lambda_{1\mu} = 0$$

$$l_2 V_2(\eta_{2i}, \alpha, \beta) + \int_{t_0}^{t_1} \sum_{\mu=1}^{m} \psi_\mu(\eta_{2i}, \alpha, \beta, t) d\lambda_{2\mu} = 0 \qquad (12)$$

$$l_r V_r(\eta_{ri}, \alpha, \beta) + \int_{t_0}^{t_1} \sum_{\mu=1}^{m} \psi_\mu(\eta_{ri}, \alpha, \beta, t) d\lambda_{r\mu} = 0$$

for every set of functions $\eta_{\rho i}$ *of class* C^1 *in* a_0, b_0 *and for all numbers* α, β.

The subscript i in the equations (12) takes on the values $1, r+1, \ldots, r+m$ in the first set, $2, r+1, \ldots, r+m$ in the second set and so on to $r, r+1, \ldots, r+m$ in the last set. There are, therefore, r sets of $m+1$ equations each, i.e., there are $r+rm$ equations in all. These $r+rm$ equations plus the m equations $\varphi_\mu = 0$ are sufficient in number to determine the rm functions of limited variation $\lambda_{\rho\mu}$ and the $r+m$ functions $y_1, y_2, \ldots, y_{r+m}$.

For the normal case we take the l_ρ all equal. Then if $W_\rho = W_{\rho+1}$, the functions of limited variation $\lambda_{\rho\mu}$, $(\rho = 1, \ldots, r)$, are solutions of the same system of Stieltjes integral equations, and, therefore, $\lambda_{\rho\mu} \equiv \lambda_{\rho+1,\mu}$ unless the l_ρ are characteristic values.[9] If the l_ρ are not characteristic values, the equations (12) for this special case reduce to the equations of Hahn as is readily seen.

* NATIONAL RESEARCH FELLOW IN MATHEMATICS.

[1] Roos, *J. Political Economy*, not yet published.

[2] Levy, *Leçons d'Analyse Fonctionelle*, pp. 50, 52.

[3] Roos, *Trans. Amer. Math. Soc.*, not yet published.

[4] Roos, these PROCEEDINGS, p. 145, March, 1927.

[5] Pareto, *Manuel d'Economie Politique*, pp. 607; Evans, these PROCEEDINGS, **11**, 90; Evans, *Amer. Math. Monthly*, 1924.

[6] See (2) or Evans, *Cambridge Colloquium*, pp. 55–66.

[7] H. L. Moore, *Quar. J. Economics*, Nov., 1926.

[8] Hahn, *Sitzungber. Akad. Wien*, 131.

[9] Roos, "On a Type of Stieltjes Integral Equation," to be offered to the *Bull. Amer. Math. Soc.*

ON THE FOUNDATIONS OF THE THEORY OF DISCONTINUOUS GROUPS OF LINEAR TRANSFORMATIONS

BY LESTER R. FORD[1]

LEIPZIG, GERMANY

Communicated March 15, 1927

1. *The Isometric Circle.*—In the present note we introduce the simple concept of the isometric circle of a linear transformation and state, largely without proofs, certain results of researches on the application of the concept to the theory of properly discontinuous groups of linear transformations.[2] The isometric circle brings to the foundations of the theory much of definiteness and simplicity. It is, furthermore, a tool of considerable utility in the study of group problems of various kinds.

Given a linear transformation

$$z' = T(z) = \frac{az + b}{cz + d}, \qquad ad - bc = 1, \qquad c \neq 0. \tag{1}$$

Then

$$dz' = \frac{dz}{(cz + d)^2}. \tag{2}$$

The complete locus of points in the neighborhood of which lengths and areas are unaltered in magnitude by the transformation; that is, for which $|dz'| = |dz|$, is the circle

$$|cz + d| = 1, \qquad \text{or} \qquad \left|z + \frac{d}{c}\right| = \frac{1}{|c|}. \tag{3}$$

This circle will be called the *isometric circle* of the transformation T. It has the center $-d/c$ and the radius $1/|c|$.

Among the properties of the isometric circle are the following:

(a_1) Lengths and areas within the isometric circle are increased in magnitude, those without the isometric circle are decreased in magnitude, when transformed by T;

(b_1) T carries the isometric circle (3) into $|cz - a| = 1$, the isometric circle of the inverse transformation T^{-1};

(c_1) T carries any point exterior to (3) into the interior of the isometric circle of T^{-1}.

2. *The Isometric Circles for the Transformations of a Group.*—Given a properly discontinuous group of linear transformations, T_1, T_2, ..., $T_i = (a_iz + b_i)/(c_iz + d_i)$. We suppose the group so transformed that there is no point congruent to infinity in a suitable neighborhood of infinity. Then $c_i \neq 0$; and the isometric circle I_i of T_i exists. We can then show that:

(a_2) The centers of the isometric circles I_i lie in a finite region (this follows from the fact that the centers, $-d_i/c_i$, are congruent to ∞);

(b_2) The centers of I_i are distinct;

(c_2) The radii of I_i are bounded;

(d_2) The number of isometric circles with radii exceeding any given positive quantity is finite.

If the group contains an infinite number of transformations the centers of I_i have one or more cluster (limit) points. Such a cluster point is called a *limit point* of the group. A point of the plane which is not a limit point is called an *ordinary point*. The transforms of a limit point are limit points; the transforms of an ordinary point are ordinary points. The following theorems can be proved:

(e_2) In the neighborhood of a limit point P there is an infinite number of distinct points congruent to any given point of the plane, with at most the exception of P itself and of one other point;

(f_2) If the set of limit points consists of more than two points it is a perfect set;

(g_2) If a closed set of points Σ, consisting of more than one point, is transformed into itself by all the transformations of the group, then Σ contains all limit points of the group.

3. *The Region R.*—Let the isometric circles of all the transformations of the group be constructed; and let R be the region exterior to all the circles. More accurately, a point z will belong to R if a circle can be drawn about z sufficiently small that it contains on its interior no point interior to an isometric circle. The region thus defined consists entirely of interior points. It contains the neighborhood of ∞. It may be a connected region or it may consist of two or more (possibly an infinite number) of disconnected parts. Concerning this region we have proved the following theorems:

(a_3) R and the regions congruent to R form a set of regions which extend into the neighborhood of every point of the plane.

(b_3) R constitutes a *fundamental region* for the group. By this is meant that no two points of R are congruent and that any region adjacent to R contains points congruent to points of R. That no two points of R are congruent follows from (c_1); for the transform of a point of R is within an isometric circle, and hence is outside R. That an adjacent region contains points congruent to points of R follows from (a_3).

(c_3) Any closed region not containing limit points of the group is covered by a finite number of transforms of R. These regions fit together without lacunae.

(d_3) Within any region enclosing a limit point of the group there lies an infinite number of transforms of the entire region R.

The preceding theorems furnish a picture of the transforms of R. R

and the regions congruent to it fit together to fill up all that part of the plane which consists of ordinary points. They cluster in infinite number about the limit points.

4. *The Boundary of R.*—In the most general case a boundary point P belongs necessarily to one of the following three categories:

(α) P is a limit point of the group;

(β) P is an ordinary point and lies on a single isometric circle;

(γ) P is an ordinary point and lies on two or more isometric circles. P is then called a *vertex*.

[It is desirable to include under (γ) the following special case. If P is a fixed point of an elliptic transformation of period two; so that, although P lies on a single isometric circle, it separates two congruent arcs on the circle, we shall classify P under (γ) rather than (β).]

It can be shown that there exist groups such that all the boundary points are of category (α).

Concerning the boundary points of category (β) we can prove the following theorem:

(a_4) The boundary points of category (β) form a set of bounding circular arcs, or *sides*, which are congruent in pairs. Two such congruent sides are equal in length.

The number of sides may be finite or infinite. A side ends in a vertex or in a limit point or consists of a whole isometric circle. The fact that congruent sides are equal in length is a consequence of the fact that the sides are arcs of isometric circles.

A boundary point of category (γ) may be congruent to other boundary points of the same category or it may not. A set of congruent vertices is called a *cycle*. The cycle may consist of one vertex or of several, but in any case the number is finite. Through each vertex of the cycle pass the same number of isometric circles. The number of vertices in the cycle cannot exceed the number of isometric circles passing through each vertex.

The sides of R which meet at a vertex make an angle which we call the *angle at the vertex*. We have the result:

(b_4) The sum of the angles at the vertices of a cycle is $2\pi/k$, where k is an integer. If $k > 1$, each vertex of the cycle is a fixed point of an elliptic transformation of period k.

5. *Generating Transformations.*—The sides of R are congruent in pairs. Let $l_1, l_1'; l_2, l_2'; \ldots$ be the pairs of congruent sides; and let S_i be the transformation of the group carrying l_i into l_i'. In many cases all transformations of the group are got by combining $S_1, S_2, \ldots$. That is, these transformations *generate* the group. The facts in the general case are covered by the following theorem:

(a_5) If every ordinary point of the plane can be joined to some point of R by a curve which does not pass through a limit point, then the trans-

formations S_1, S_2, ... constitute a set of generating transformations for the group; otherwise they do not.

[1] NATIONAL RESEARCH FELLOW IN MATHEMATICS.

[2] Certain results of this note in the special case of Fuchsian groups have been published. L. R. Ford, "The Fundamental Region for a Fuchsian Group," *Bull. Amer. Math. Soc.*, **31**, 531–539 (1925).

ON THE FORMATION OF GROUPS OF LINEAR TRANSFORMATIONS BY COMBINATION

BY LESTER R. FORD[1]

LEIPZIG, GERMANY

Communicated March 15, 1927

In this note we discuss groups formed by combining a finite or infinite number of known groups. By suitable combinations we can set up a great variety of Kleinian groups. Our present knowledge of Kleinian groups is not extensive. These groups in all their generality present a complexity and a richness which are largely unfathomed. By the method here described we can form broad classes of groups which bring out certain of the intricate possibilities.

In a previous note we gave a method of constructing a fundamental region for a group.[2] The fundamental region consists of all those points z such that in the neighborhood of z there are no points interior to an isometric circle of a transformation of the group. We consider a set of groups Γ_1, Γ_2, ..., finite or infinite in number, such that the isometric circles of the transformations of each group are exterior to (external tangency will be permitted) the isometric circles of the transformations of all the other groups.

Let the transformations of Γ_1, Γ_2, ... be combined in all possible ways to form a group Γ. We have two kinds of transformations: (1) Those belonging to the original groups; and (2) *cross products* resulting from the combination of transformations not all belonging to the same original group. Our major result is the following:

THEOREM.—*The isometric circle of each cross product lies within the isometric circle of some transformation belonging to one of the original groups.*

It follows from this theorem that in the formation of the fundamental region for Γ we can disregard the isometric circles of all cross products. The fundamental region R for Γ consists of all points common to the fundamental regions R_1, R_2, ... of Γ_1, Γ_2, ..., exclusive of those common points (if any) which have an infinite number of boundary points of R_1, R_2, ..., in their neighborhoods.

The proof of the theorem depends upon the following lemma. We represent the isometric circle of a transformation T by I_t and the isometric circle of its universe by I'_t. We use the symbol $<$ to mean "is contained within."

LEMMA.—*If I_t and I'_s are exterior to one another then for the transformation*

$$U = TS \text{ we have } I_u < I_s.$$

The proof is simple. Let z be a point outside (or on) I_s. Then S carries z into z' within (or on) I'_s with decrease of lengths (or without alteration of lengths) in the neighborhood[3] of z. Now z' is outside I_t; so T transforms z' with decrease of lengths. Hence U transforms z with decrease of lengths, whence z is outside I_u. Since every point on or outside I_s is outside I_u, the latter circle is contained in the former.

Let $T^{(i)}_{m1}$, $T^{(i)}_{m2}$, ... be the transformations of Γ_i. As a convenient notation let the subscripts be so assigned that $T^{(i)}_{-n}$ is the inverse of $T^{(i)}_{n}$. [There are thus two symbols for a self-inverse transformation, but this leads to no confusion.] The group Γ consists of all transformations of the form

$$T = T^{(n_k)}_{m_k} T^{(n_{k-1})}_{m_{k-1}} \cdots T^{(n_2)}_{m_2} T^{(n_1)}_{m_1}.$$

We can suppose that no factor is the identical transformation, since it can be cancelled; also that no two successive indices n_i, n_{i+1} are the same, for the two transformations can be combined into a single transformation of the group Γ_{n_i}. We shall represent the isometric circle of T by $I^{(n_k, \ldots, n_1)}_{m_k, \ldots, m_1}$

We shall prove that

$$I^{(n_k, n_{k-1}, \ldots, n_1)}_{m_k, m_{k-1}, \ldots, m_1} < I^{(n_{k-1}, \ldots, n_1)}_{m_{k-1}, \ldots, m_1} < \ldots < I^{(n_2, n_1)}_{m_2, m_1} < I^{(n_1)}_{m_1}. \qquad (a)$$

The inequality holds for all combinations of two transformations

$$I^{(n_2, n_1)}_{m_2, m_1} < I^{(n_1)}_{m_1},$$

for $I^{(n_2)}_{m_2}$ and $I^{(n_1)}_{-m_1}$ are, by hypothesis, exterior to one another, and the inequality follows from the lemma.

We prove the general case by induction. Supposing the inequalities to hold for all combinations of $k-1$ transformations, we prove them for k transformations. In (a) all the inequalities except the first hold. Put

$$S = T^{(n_{k-1})}_{m_{k-1}} \cdots T^{(n_1)}_{m_1}; \qquad \text{then} \qquad S^{-1} = T^{(n_1)}_{-m_1} \cdots T^{(n_{k-1})}_{-m_{k-1}}.$$

Now S^{-1} is a combination of $k-1$ transformations; hence its isometric circle satisfies the inequality $I'_s < I^{(n_{k-1})}_{-m_{k-1}}$. Now $I^{(n_{k-1})}_{-m_{k-1}}$ and $I^{(n_k)}_{m_k}$ are exterior to one another; hence $I^{(n_k)}_{m_k}$ and I'_s are exterior to one another. Hence applying the lemma to the transformation $T = T^{(n_k)}_{m_k} S$ we have that

$$I_t < I_s,$$

which is the first inequality of (a). This establishes (a) for all cases.

The proof of the theorem stated above follows at once. For each cross product we can form a chain of inequalities (a). The last circle in (a), which includes all the preceding, is the isometric circle of a transformation of one of the original groups.

We make an application of the method. The simplest group consists of the identical transformation and a single self-inverse transformation. The latter transformation can be so chosen that any given circle is its isometric circle. The fundamental region for this group is the exterior of the given circle. Let C be a closed curve. We can fill C with circles $I_1, I_2, \ldots$ such that every point within C is within a circle or has circles in its neighborhood. Let T_i be the self-inverse transformation with I_i as isometric circle. Then the group Γ generated by $T_1, T_2, \ldots$ has as fundamental region the exterior of C; the boundary of the region (C itself) consists entirely of limit points; and there is a complicated set of limit points lying within C.

We get instructive groups by combining a finite or infinite number of cyclic groups, finite groups, Fuchsian groups, transformed doubly periodic groups, and the like.

[1] NATIONAL RESEARCH FELLOW IN MATHEMATICS.

[2] "On the Foundations of the Theory of Discontinuous Groups of Linear Transformation," these PROCEEDINGS, **13**, 286 (1927).

[3] Loc. cit. (a_1), (c_1).

A STRATIGRAPHIC STUDY IN THE GILA-SALT REGION, ARIZONA

BY ERICH F. SCHMIDT

DEPARTMENT OF ANTHROPOLOGY, COLUMBIA UNIVERSITY

Communicated March 22, 1927

The present study is the extract of a report soon to be published on "The Time Relations of Prehistoric Pottery Types in Southern Arizona."

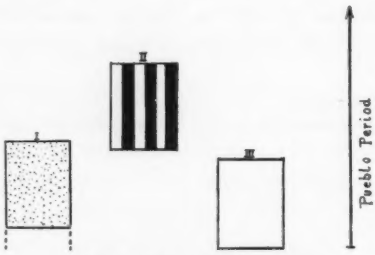
The methods of excavation employed in the Mrs. W. B. Thompson Expedition, which was carried on in 1925 under the auspices of the American Museum of Natural History, were determined by the all-important question as to the chronological relations of the prehistoric remains encountered during the work.

It is now a generally recognized fact that no other product of prehistoric workmanship records cultural changes and relations more sensitively than

pottery, and for that reason it furnishes an instrument of great help to the student of cultural development in times without written records. The manner in which this instrument was used in establishing a basis for such a study in the area under consideration is presented herein.

The work centered principally around three localities and their environment—the ruins Togetzoge, Spring Creek and Pueblo Grande. The ruin Togetzoge, a one-story pueblo of the "checker-board" type, is situated in the mountains between Superior and Miami opposite to Irion's Ranch. It comprises about 120 rooms, built of unworked boulders and is enclosed by a defense wall. About one-third of this settlement was excavated and the contents of the common burial ground recovered. Here as well as in ruins east of Togetzoge, near Globe, Cutter and Rice, the majority of the decorated pottery is "Gila polychrome" with a small admixture of certain foreign wares.

In ruin Spring Creek, a two-story pueblo which lies three miles north of Spring Creek store near the southern shore of Roosevelt Lake, only two rooms were excavated. But six shallow refuse mounds associated with this pueblo were closely investigated. The depth of these deposits was not sufficient to yield any stratigraphic results. But the "pottery scale" was established indicating that here "black-on-white" of a hybrid Upper Gila and local type prevails. This black-on-white was tentatively included in the Little Colorado group of ceramic types which was also represented by some types of red ware, frequently encountered in the north but occurring here only in small numbers. Gila polychrome was almost entirely absent in the refuse heaps, though fragments of this type were found on the surface and in the upper rooms of the building. The home of a third principal pottery type—"red-on-yellow," will now be considered. Sherds of this ware occur at Spring Creek and less frequently also at Togetzoge.



I—Lower Salt red-on-yellow
II—Central Gila polychrome
III—Little Colorado black-on-white

There is an extensive ruin popularly called Pueblo Grande, about seven miles east of Phoenix. It is probably three stories high, the walls being built of alternating flat stone slabs and adobe. Here again an enclosure surrounds the building, which is not far from the junction of an ancient irrigation canal with the Salt River. This ruin, one of many existing in the vicinity of Phoenix, is of particular interest, since with it is associated a good-sized refuse heap, giving an opportunity for stratigraphic research.

The sectioning was carried on after cleaning the surface of the apex by

cutting an area of 7.50 square m. in successive layers 15 cm. deep. The potsherds occurring in the various strata were distributed as follows:

	DECORATED											UNDECORATED		
	LOWER SALT			CENTRAL GILA	LITTLE COLORADO					UNKNOWN ORIGIN				
	RED-ON-YELLOW	INCISED[1] RED-ON-YELLOW		POLYCHROME	BLACK-ON-WHITE	WHITE BORDERED[2] BLACK-ON-RED	BLACK-ON-BUFF	SLIGHTLY INDENTED		RED BROWN-ON-BUFF BLACK-POLISHED-INTERIOR		BLACK-POLISHED-INTERIOR	PLAIN	
I	2		2	6						1	9	195	1541	1745
II	3		3	2							5	204	1052	1261
III				4		2	1	1	4	1	9	252	998	1259
IV	9	1	10								10	258	939	1207
V	5		5		1				1		6	184	617	807
VI	4	1	5								5	159	695	859
VII	14	1	15	1							16	149	847	1012
VIII	9		9		1				1		10	207	788	1005
IX	17	1	18			1			1		19	212	985	1216
X	17	2	19								19	233	1124	1376
XI	23		23	1							24	243	1109	1376
XII	20	5	25	2							27	200	1254	1481
XIII	28	6	34		1				1		35	208	1163	1406
XIV	25	11	36								36	161	958	1155
XV	13	10	23								23	59	497	579
XVI	12	1	13		1				1		14	25	212	251
XVII	5	1	6								6	10	86	102
XVIII	9	1	10								10	8	59	77
	215	41	256	16	4	3	1	1	9	2	283	2967	14924	18174

[1] Distinguished from red-on-yellow solely by additional ornamentation of incised lines on the exterior, parallel to the rim.

[2] Fine ware with whitish paste and red slip, the decorations in black bordered by thin white lines (also dentiform patterns in white).

The proportions of the principal pottery classes are as follows:

Undecorated	Plain	14924	81.1%
	Black-polished-interior	2967	16.2%
Decorated	Painted	282	1.6%
	Slightly indented	1	

The most striking pecularities are:

1. The total absence of corrugated-indented ware, abundant in most areas of the Southwest. (The single "slightly indented" sherd is a drift piece from the Roosevelt Lake district.)

2. The exceedingly small percentage of painted ware.

Considering now the distribution of the total of sherds and that of the main groups as represented in graph, figure 2, we notice two maxima in the total curve (*d*) in strata XII and I, with a considerable decrease of sherds in between, with a minimum in stratum V. The curve of the decorated ware (*a*) does not ascend to a second maximum as the total curve, but remains stable in the upper strata after a steady decline upward from the apex in strata XIII and XIV. These phenomena will be considered below.

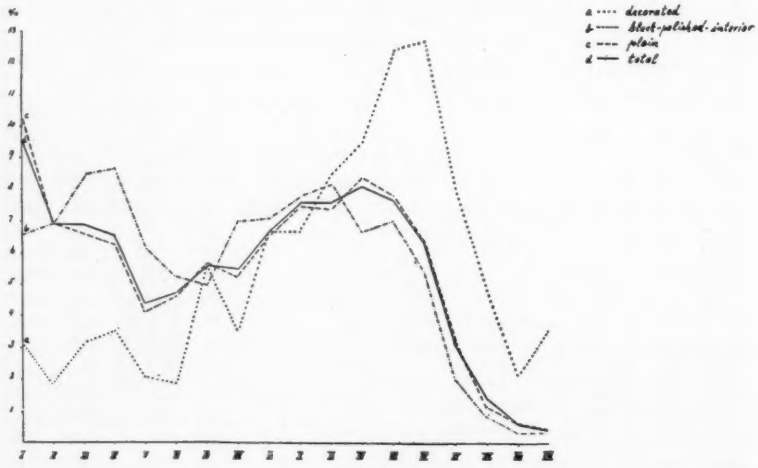


FIGURE 2

	a	*b*	*c*	*d*		*a*	*b*	*c*	*d*
I	3.2	6.6	10.3	9.6	X	6.7	7.8	7.5	7.6
II	1.8	6.9	7.0	6.9	XI	8.5	8.2	7.4	7.6
III	3.2	8.5	6.7	6.9	XII	9.5	6.7	8.4	8.1
IV	3.5	8.7	6.3	6.6	XIII	12.4	7.0	7.8	7.7
V	2.1	6.2	4.1	4.4	XIV	12.7	5.4	6.4	6.3
VI	1.8	5.3	4.6	4.7	XV	8.1	2.0	3.3	3.2
VII	5.6	5.0	5.7	5.6	XVI	4.9	0.8	1.2	1.4
VIII	3.5	7.0	5.3	5.5	XVII	2.1	0.3	0.6	0.6
IX	6.7	7.1	6.6	6.7	XVIII	3.5	0.3	0.4	0.4
						100%	100%	100%	100%

Graph, figure 3, shows the distribution of the decorated types, i.e., the units which compose curve *a*, figure 2. The Little Colorado group will be further subdivided.

Following are the most prominent features illustrated by graph, figure 3:

1. The steady increase of Lower Salt ware (curve *a*) from the top of the refuse mound toward the bottom and the striking predominance of

this ware in the lower strata. These two facts show beyond doubt that Pueblo Grande was built by the makers of this pottery.

2. While sherds of foreign types occur below stratum III only sporadically, there is an intermingling of all wares in the upper three strata, with a prevalence of Central Gila polychrome (curve *b*). This predominance of the latter ware in the top strata in contrast to the prevalence of Lower Salt in the remaining layers and the disappearance of stray pieces of the Gila ware above the strata of the maximal occurrence of Lower Salt permit

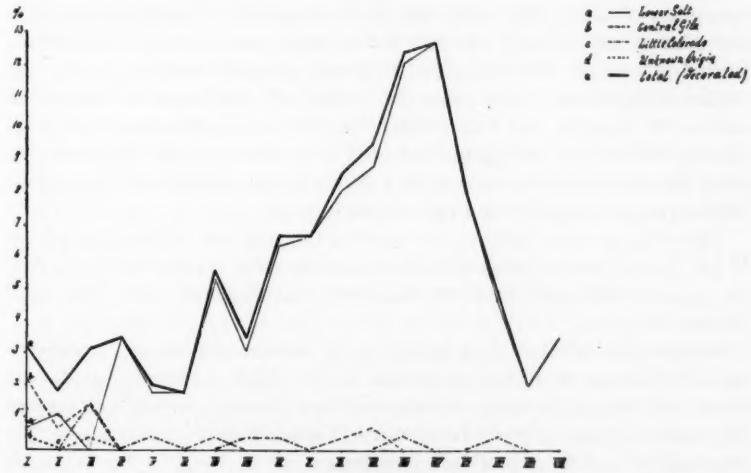


FIGURE 3

	a	*b*	*c*	*d*		*a*	*b*	*c*	*d*
I	0.7	2.2		0.35	X	6.75			
II	1.1	0.7			XI	8.1	0.35		
III		1.4	1.4	0.35	XII	8.8	0.7		
IV	3.5				XIII	12.05		0.35	
V	1.75		0.35		XIV	12.7			
VI	1.75				XV	8.1			
VII	5.25	0.35			XVI	4.6		0.35	
VIII	3.15		0.35		XVII	2.1			
IX	6.4		0.35		XVIII	3.5			

100%

the conclusion that the Lower Salt ware is considerably older than (the) Central Gila polychrome. The latter type appeared when the manufacture of the locally indigenous ware approached its end, or even after it had ceased to be made at Pueblo Grande. Here attention must be directed to the coincidence between the principal occurrence of Gila sherds and the second increase of the total of sherds, i.e., mainly plain ware, in the upper

strata. A similar correlation has been noticed at a second locality of the same region. (See below, "La Ciudad.")

In view of the importance of the problem as to the chronological relations of the two decorated types, and considering the small number of Gila sherds which had to serve as a basis for far-reaching conclusions, the results were checked up by a second stratigraphic study. In the refuse heap of La Ciudad, a ruin situated within the city limits of Phoenix, the conditions almost exactly paralleled those encountered at Pueblo Grande. Gila polychrome occurred solely in the first five strata out of fourteen. The maximum of Lower Salt ware was in stratum VII. Further proof as to the relative antiquity of the two wares was obtained in the domiciliary mound associated with the refuse heap and certainly built up during successive occupations. In a room, the outlines of which were visible on the surface of the apex, two fragmentary vessels were found, one a Gila polychrome, the other, corrugated-indented.[1] Underneath the floors of this room the trench which was sunk to a depth of 4 m. struck only Lower Salt sherds and one fragment of black-on-white.

The evidence furnished by the work in the ruin and in the refuse heap of La Ciudad proves in a satisfactory way that the conclusion concerning the greater antiquity of Lower Salt ware as compared with Gila polychrome is correct.

Returning to Pueblo Grande, one more question has to be considered: the distribution of sherds comprised in the Little Colorado group. It would take too much space to elaborate here the evidences which concern the time relations of the various Little Colorado types with Lower Salt ware on the one hand and Gila polychrome on the other. The relations of some types are not known at all, principally because we know so little about the situation in the southern part of the Little Colorado area.[2]

As to the few sherds belonging to Little Colorado types occurring in the refuse mound of Pueblo Grande, there is an accumulation of all types in stratum III, except one: black-on-white. The sherds of this type are distributed over the remaining strata. This fact is significant only if we draw other criteria into consideration. At the polychrome ruin, Togetzoge, among thousands of sherds not even one black-on-white fragment was found, though the majority of foreign sherds came from the Little Colorado region. They were principally black-on-red, white bordered black-on-red, and a few sherds of black(glaze)-on-brown. How shall we explain this, since out of all the Little Colorado settlements the black-on-white ruins of Roosevelt Lake are situated nearest to Togetzoge, the red ware types representing the principal intrusive pottery at Togetzoge constituting only a small percentage of the decorated pottery in the Roosevelt Lake sites? The answer is not difficult. Stratigraphic work pursued by L. Spier and F. W. Hodge in the northern part of the Little Colo-

rado region showed that black-on-white preceded all the other types found in that territory; black-on-red appearing while it declined. The situation in the south of this area seems to be approximately the same: After a decline of black-on-white, types of red ware were made which found their way to the Pueblo Togetzoge, not existing during the black-on-white period.

As to the chronological relations of black-on-white and Lower Salt, they are synchronous, but to what extent is not known. The principal clue for their contemporaneity was found in a burial ground near Armer's Gulch on the northern shore of Roosevelt Lake. Here fragmentary vessels of red-on-yellow and black-on-white were found in the same grave. As further criterion might be mentioned the relative frequency of Lower Salt sherds in the refuse mounds at Spring Creek and in the other black-on-white ruins of the Roosevelt Lake district.

Summary.—The general pottery survey in addition to the stratigraphic study near Phoenix showed that there exist within the boundaries of the Lower Gila region three ceramic areas, characterized by three principal decorated types:

In the east: the Central Gila area with Gila polychrome.

In the north: The Roosevelt Lake district with black-on-white.

In the west: the Lower Salt area with red-on-yellow.

In regard to the latter area, it was emphasized that the makers of red-on-yellow never made corrugated-indented ware, in contrast to the manufacturers of the two other types, especially those of black-on-white.

The chronological relations of the three main types, as explained on the preceding pages, are best illustrated by the following schematic diagram:

In the territory under consideration, black-on-white and red-on-yellow fall in the beginning of the Pueblo period. Gila polychrome follows later and is probably synchronous with the last phase of the culture characterized by red-on-yellow.

A great number of most interesting problems arise now and others, vague before, appear more distinct. A comparative study of the three culture complexes, defined by the respective pottery types will have to deal with all of them without neglecting, of course, the omnipresent questions as to the origin and as to the possible affinity of the pre-Spanish cultures with those of the historic tribes—the Yumans and the Pimans being for us of particular interest.

As a striking example of the application of the results gained by the stratigraphic work one case is here mentioned. During his excavation of "Los Muertos," situated in the Lower Salt region, T. H. Cushing encountered there two methods of disposal of the dead, inhumation and cremation. He drew the conclusion that the different treatment of the dead was due to a difference in their social position, the "priestly class" having

been interred. But an examination of the photographs taken by him and kept in the Peabody Museum, Harvard University, shows that the vessels containing cremated remains are of the red-on-yellow type; while those accompanying the interred skeletons are Gila polychrome.

The duality of the methods of disposal was therefore not due to a difference in social position but to a difference in time. The original builders of Los Muertos cremated their dead while during a later period, most probably during a second occupation by a different group of people, inhumation was introduced.

[1] Corrugated-indented was entirely lacking in the refuse mound of La Ciudad as it was at Pueblo Grande.

[2] The same is true for the "Unknown ware" (table, Fig. 1). The home of this type is not yet known, but vessels and sherds were found at Togetzoge.

THE INFRA-RED ABSORPTION SPECTRA OF ACETYLENE (C_2H_2), ETHYLENE (C_2H_4) AND ETHANE (C_2H_6)

By Charles F. Meyer and Aaron A. Levin

Department of Physics, University of Michigan

Communicated April 11, 1927

All bands of the vibrational-rotational type which have hitherto been resolved into lines, originate from molecules which contain hydrogen and only one atom of another kind. Molecules containing two or more other atoms have, in general, a much higher moment of inertia, and investigators have felt that there might be extreme difficulty in resolving the bands due to them or that they would prove too complex to admit of analysis. The present investigation of the spectra of acetylene, ethylene and ethane reveals fine structure of the bands of these compounds which each contain two atoms of carbon. The structure is of an interesting nature and in some cases simple. Moreover, there are regularities of a type not previously observed in vibrational-rotational bands.

Various regions of absorption between 3.μ and 15.μ have been investigated. The data in the region of longest wave-lengths are nearly complete, and the authors feel that the publication of the curves will be of interest, especially as some months will elapse before the entire investigation is concluded.

A grating spectrometer is being used which is of a type described by others.[1] There are modifications in some rather important details, but these will not be discussed. The thermoelement was constructed according to the method of Dr. A. H. Pfund[2] and under his kind supervision.

It is provided with a calcium evacuator.[3] The grating is one of 1440 lines per inch ruled by Dr. E. F. Barker on the ruling machine designed by him and Mr. A. de Khotinsky.

In figure 1 are shown curves obtained by Dr. W. W. Coblentz[4] with a prism spectrometer. These represent the absorption of the gases under

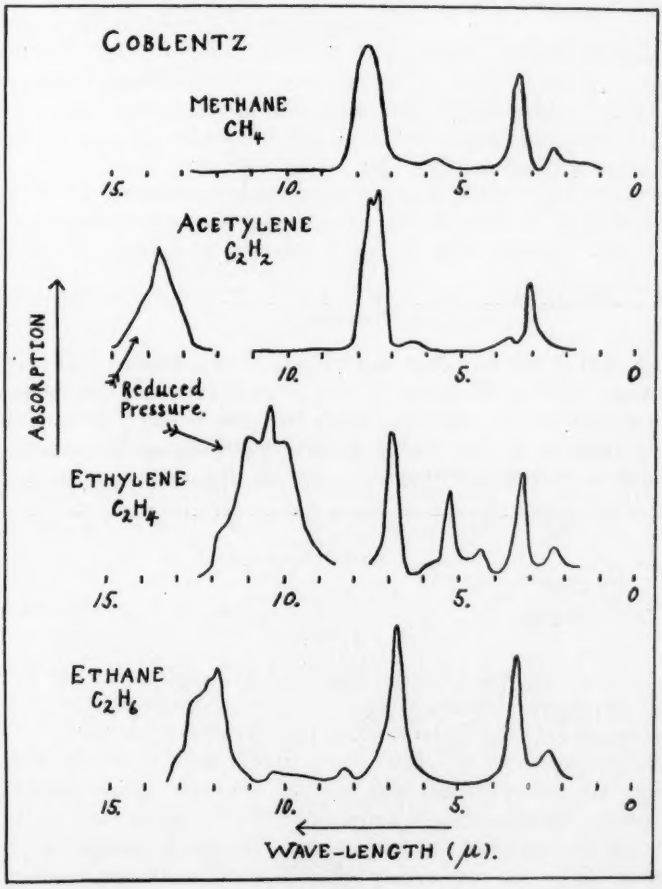


FIGURE 1

investigation and also that of methane, at the top, for comparison. The curves have been re-drawn so as to show absorption increasing upwards.

In the present work it seemed desirable to begin with an examination of the absorption region of greatest wave-length for each gas, as this

probably represents one of the fundamental frequencies of vibration. For ethylene this region lies at a shorter wave-length than for the other gases, and is hence the most accessible to measurement. Accordingly, it was first investigated. A typical curve obtained is shown in figure 2. The intense central maximum at 10.53μ doubtlessly represents a "Q"

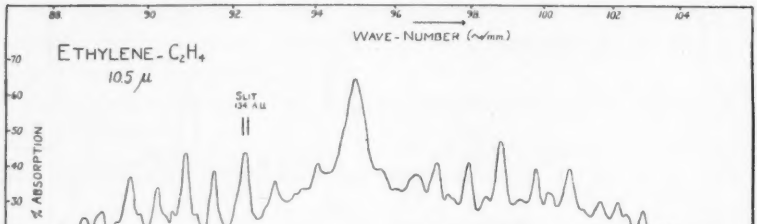


FIGURE 2

branch, and it will be noted that Coblentz' curve showed evidence of its existence. The region was gone over several times and the outstanding characteristics of the curve may safely be taken as real. That is, the intensity relations for the lateral maxima which appear in the figure are repeated in various measurements, and the fainter lines lying between

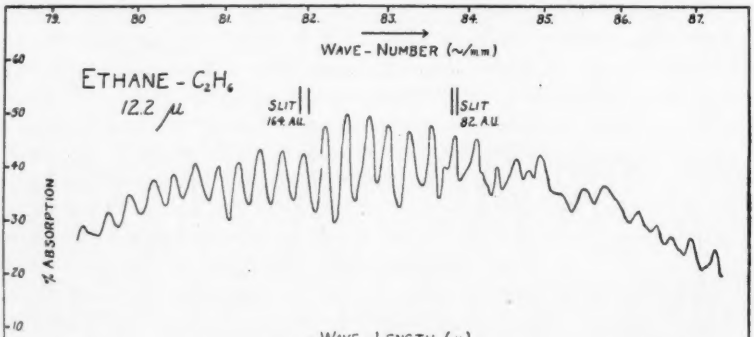


FIGURE 3

the more intense ones are recurrent. Interpretation of this band and others will not be presented in this paper.

In figure 3 is shown a curve for ethane. It will be noted that there is a succession of maxima which is regular both in regard to variation of

intensity and of spacing. Because of the decrease of spacing in proceeding toward shorter wave-lengths it was thought desirable to narrow the slit at about the middle of the band. The resulting improvement in definition is made evident by a rise at the maxima of absorption and by a depression at the minima. From 11.9μ to 11.6μ it will be noticed that the curve is irregular. Repeated observation leaves the question of reality of these irregularities still unanswered.

Figure 4 shows the curve obtained for acetylene. There is again an intense central maximum or "Q" branch. In the "R" branch, on the right, there is a well-marked alternation of intensity of the lines. This is the first time such an alternation has been observed in a vibrational-rotational band, except that the ethylene band, figure 2, also gives some indication of it. On the extreme right of figure 4 is a detached curve which was taken with more gas in the cell. In the "P" branch, on the left of the figure, the alternation of intensity is not so well marked. The sensi-

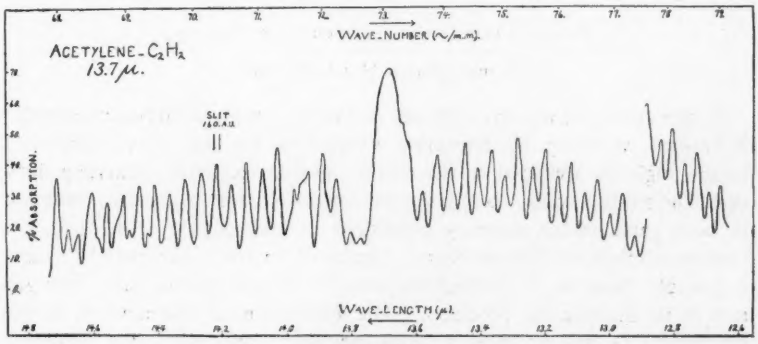


FIGURE 4

tivity of the instrument falls off rapidly toward greater wave-lengths and there is consequently a great increase in observational error. The failure to show pronouncedly the alternation of intensity and the occurrence of irregularities from 13.9μ toward greater wave-length are attributed to observational error. We have thus far only one curve for this region. The intense lines of the "P" branch form a regular sequence of spacing with the faint ones of the "R" branch, and the faint ones of the "P" branch form a regular sequence with the intense ones of the "R" branch.

The absorption band of acetylene at 7.5μ which is shown in Coblentz' curves, see figure 1, has been examined with a grating of 2880 lines per inch ruled on the first Rowland ruling machine. The "Q" branch is entirely absent. The alternation of intensity is shown to a marked degree throughout both the "P" and "R" branches. Moreover, the sequences formed by the intense lines of each branch with the faint lines of the

other can be readily followed by simple inspection. There are no missing lines at the center of the band.

Preliminary data have been obtained in five other regions of absorption shown in Coblentz' curves. In some of these the measurements have progressed quite far but they will not be discussed at present, for it is desired to repeat them and possibly increase their refinement by improvements in some parts of the apparatus.

[1] E.g., Sleator, W. W., *Astrophys. J.*, **48**, 1918 (125–143).
[2] *Physik. Zeitschr.*, **12**, 1912 (870–873).
[3] Coblentz, W. W., *J. O. S. A.*, **5**, 1921 (356–362).
[4] *Ibid.*, Publications of the Carnegie Institution of Washington, **35**, 1905 (166–169).

THE SPECTRUM OF CO_2

BY G. W. FOX, O. S. DUFFENDACK AND E. F. BARKER

PHYSICS LABORATORY, UNIVERSITY OF MICHIGAN

Communicated March 28, 1927

In the course of an investigation of the spectrum of carbon monoxide[1] it became necessary to determine whether or not any of the observed bands might be assigned to the carbon dioxide molecule. Existing data upon the subject cannot be considered conclusive, for the spectra observed in both gases under ordinary conditions of excitation are identical. A continuous flow method previously employed for the study of HCl[2] makes it possible, however, to distinguish between the spectra due to a given gas and to its dissociation products. The application of this method to the study of CO_2 has revealed a band spectrum apparently not observed previously which may be ascribed with confidence to the CO_2 molecule. A few of the most intense heads did appear rather faintly on the earlier CO plates; hence it appears that in the electrical discharge through stagnant gas the reaction $2CO \longrightarrow CO_2 + C$ occurs.

A supply of CO_2 free from CO was obtained by dropping H_2SO_4 upon $CaCO_3$ in an evacuated flask (Fig. 1). Both reagents were as pure as could be obtained. The gas was collected in a trap immersed in liquid air, pumping being continued meanwhile. After the generator was shut off a fraction of this condensed gas was allowed to expand into a reservoir where it was stored at a pressure of about two atmospheres: the remainder was released through the generator. From the reservoir CO_2 was admitted slowly into the discharge tube and subjected to electron bombardment in a force free space before reaching the filament. The tube is provided with a tightly fitting diaphragm having an aperture about 2×6 mm. in front of the filament, the aperture being covered with a fine platinum gauze.

The exciting electrons are accelerated between the filament and the gauze in a direction opposite to the flow of the gas. When the proper rate of pumping is maintained—the pressure in the tube being about 0.1 mm.—there is no indication of carbon monoxide appearing in the excitation chamber, either by diffusion from the region around the filament or by dissociation of CO_2 through electron impacts. The spectra, photographed through a quartz window at one side of the tube, show none of the bands ordinarily assigned to CO. As soon as pumping ceases and the flow is stopped, however, the character of the discharge rapidly changes and the CO_2 bands disappear. This effect is observed even with an oxide-coated filament operating at a comparatively low temperature. The discharge in the

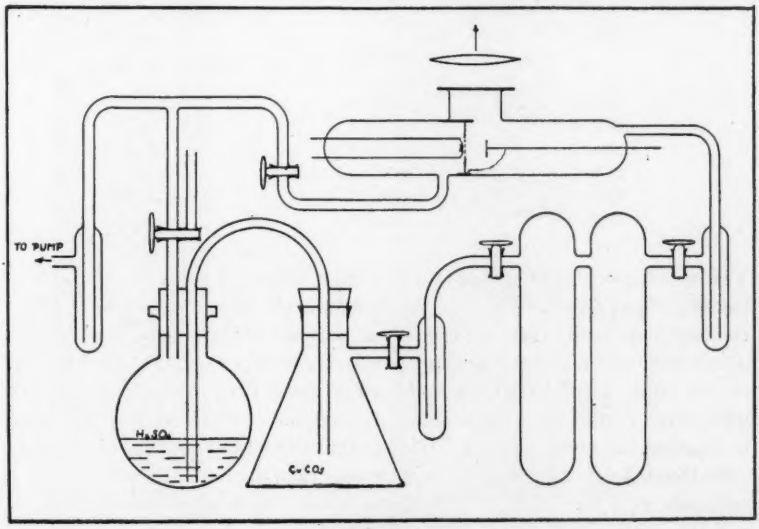


FIGURE 1
Diagram of apparatus.

flowing gas produces a violet glow between the gauze and plate which changes rapidly to the familiar sky blue color characteristic of CO when the flow is stopped. Further evidence that no dissociation products found their way into the impact region lies in the fact that the 2537 mercury line appears only very faintly on the spectrograms, even though no liquid air was applied to the trap between the discharge tube and pump. In fact the line as observed would be expected because of the mercury vapor which was known to exist in the CO_2 storage reservoirs.

Reproductions of the spectrum of the discharge through flowing CO_2 are shown in plate I. The spectrogram *A* was taken with a Hilger E2

quartz spectrograph while *B* was taken with the E1 model. The spectrum consists of bands of several different types, as revealed in the reproductions, and extends from approximately 5000 Å to 2800 Å, the bands at the ends being too faint to reproduce. For the most part, the bands are degraded toward the red and have sharp edges. The fine structure is not resolved near the edges by the E1 spectrograph nor is it revealed by a three-fold enlargement of the spectrogram in which a dispersion of approximately 1 Å per mm. is attained.

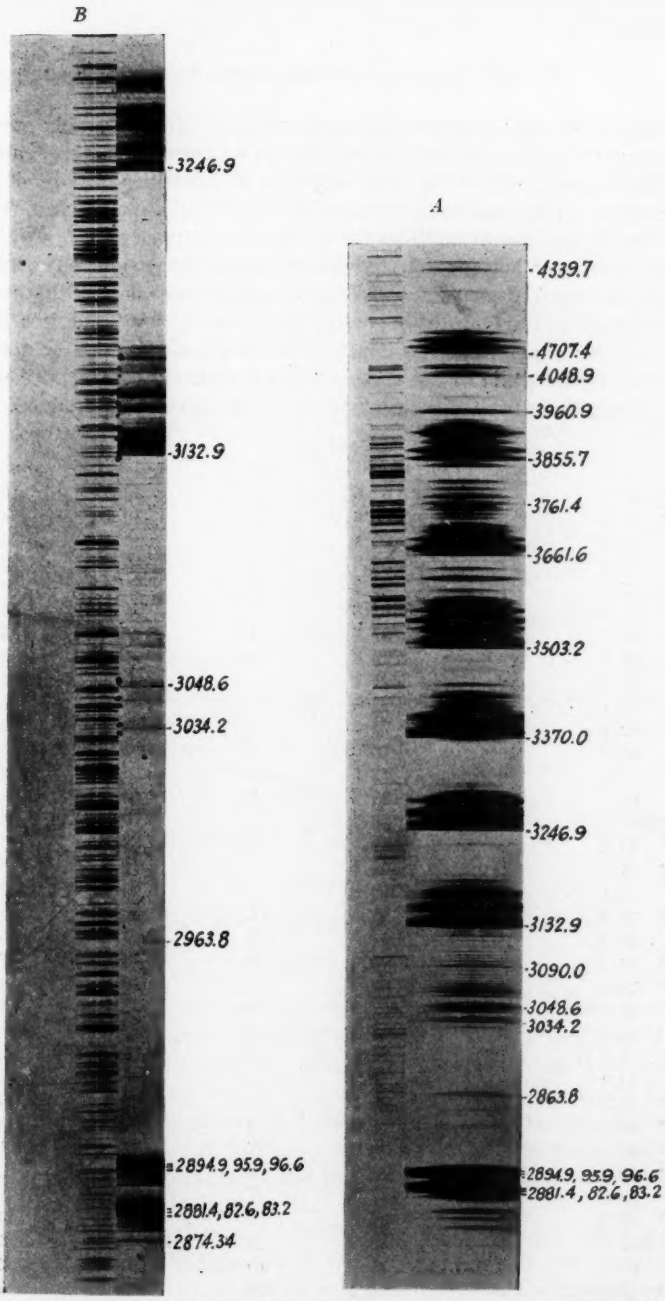
The bands constituting the strong groups consist of two or more band systems which overlap one another to some extent. Some progress has been made in the analysis of these groups and it seems certain that one system consists of the bands listed in table 1.

TABLE 1

λ	ν (VAC.)	Δν
3027.0	33026.3	
3132.9	31910.1	1116.2
3246.9	30789.8	1120.3
3370.0	29665.1	1124.7
3503.2	28537.2	1127.9
3647.8	27406.0	1131.2

This system probably extends farther than indicated in the table both in the direction of greater and in the direction of shorter wave-lengths, but the bands are fainter and are difficult to pick out of the overlapping groups. The nature of the groups associated with these edges suggests that several of the edges involve the same electronic transitions, but the frequency differences within the groups are quite irregular. We have not been able to analyze the groups on the basis of transitions between sets of multiple vibrational levels. It seems fairly certain that the bands listed in table 1 originate in a set of initial vibrational levels, the final state consisting of a single vibrational level. The related edges in the groups may result from multiple electronic terms or from a structure of the molecule such that more than one moment of inertia is involved.

The bands with edges near 2880 and 2895 are distinctly different from the others in structure. The spectrogram *A* shows two intensity maxima in this region which are resolved by the larger spectrograph into double-edged bands as shown in spectrogram *B*. Each band has two oppositely degraded branches with sharp inner edges between which a very narrow band appears. These oppositely degraded bands have some resemblance to the ammonia bands at 3360 and 3370 Å, but in ammonia the narrow band between the edges does not appear. This narrow band may be due to a zero branch which is not extensively developed. No relation between these bands and those constituting the groups has been discovered. They



Spectrogram *A*. The spectrum of flowing CO_2 taken with a Hilger E2 quartz spectrograph. Iron comparison spectrum. Spectrogram *B*. A portion of the spectrum of flowing CO_2 taken with a Hilger E1 quartz spectrograph. Iron comparison.

will have to be photographed with an instrument of considerably greater dispersion before any definite conclusions can be drawn as to their nature.

Table 2 gives a list of the wave-lengths and intensities of most of the band edges. Those marked * were measured only on a spectrogram taken with the E2 instrument and may be in error by as much as 1 Å as it was difficult to locate the edge exactly with such small dispersion. The wave-lengths not marked were measured on spectrograms taken with the E1 spectrograph and are probably correct to at least 0.1 Å. The iron lines were used for comparison. As none of these bands are absorbed in normal CO_2, the normal state of the molecule is not their final state. Whether they belong to an excited molecule or to an ion was not determined.

TABLE 2

λ	INTEN.	λ	INTEN.	λ	INTEN.
*2849.5	4	64.6	9	40.3	5
* 56.4	5	69.9	10	49.1	4
* 62.5	6	80.5	3	49.7	4
74.3	7	83.3	4	56.4	2
77.2	5	84.3	8	61.4	6
78.2	5	85.8	3	74.6	6
80.7	6	87.7	3	89.3	5
82.1	10	92.6	4	3805.4	5
82.6	9	3301.1	4	09.8	5
83.1	10	70.0	10	38.8	6
90.4	7	72.5	5	53.2	4
95.4	10	75.2	3	53.8	4
95.9	9	77.5	10	55.7	9
96.4	10	82.2	5	66.8	3
*2907.5	0	86.2	3	70.5	5
* 10.2	0	88.9	7	72.7	7
* 16.3	0	94.5	7	74.5	6
* 19.1	0	3396.9	3	90.4	5
* 24.8	1	99.4	4	99.3	5
* 25.2	1	99.9	3	3904.5	3
* 28.3	1	3400.9	3	14.1	7
* 29.4	1	02.2	1	20.2	4
* 33.1	3	04.3	1	22.2	4
* 35.7	3	13.3	3	27.3	4
* 45.2	2	26.6	3	59.4	3
* 49.4	4	43.5	2	60.9	7
* 56.8	0	* 64.5	0	64.1	5
* 58.7	2	* 71.0	0	*4002.3	2
* 61.8	2	3503.2	7	* 25.0	1
* 63.8	5	10.8	9	48.9	3
* 88.6	2	17.1	7	50.5	5
* 90.3	2	28.3	4	58.6	4
*3005.3	3	33.8	9	68.1	5
* 07.3	3	45.9	10	70.7	5
* 17.8	2	51.4	7	71.7	4
* 19.6	0	62.2	9	95.7	3

* 22.6	1	65.5	6	4103.4	3
27.0	3	81.1	4	07.9	7
34.2	6	83.3	2	09.2	5
43.9	5	86.0	2	20.8	6
48.6	6	* 99.1	6	22.6	8
* 58.3	5	3615.8	3	37.6	7
* 63.5	5	16.9	3	40.1	3
* 67.7	5	18.6	3	44.7	2
* 72.1	4	21.0	7	59.5	4
* 90.0	5	34.3	2	70.5	2
3132.9	8	37.7	6	* 85.3	2
34.6	8	38.5	6	* 96.6	0
35.7	6	47.8	3	*4208.6	1
36.7	9	61.6	7	* 64.0	3
39.2	9	68.1	6	* 81.6	4
42.4	5	70.1	6	4308.4	3
49.5	6	74.1	9	23.8	4
50.7	6	79.6	7	39.7	6
55.2	9	83.1	2	41.0	5
64.9	5	87.9	2	4359.6	5
70.1	7	3690.9	4	64.7	3
74.3	5	91.8	8	67.4	3
3246.9	10	3708.0	3		
53.9	10	13.9	3		

[1] Duffendack and Fox, *Nature*, **118**, 12 (1926); *Science*, **64**, 277 (1926); *Astrophys. J.*, in press.

[2] Barker and Duffendack, *Phys. Rev.*, **26**, 339 (1925).

THE ENTROPY OF RADIATION

By Gilbert N. Lewis

Department of Chemistry, University of California

Communicated April 6, 1927

The very close agreement between the formula of Planck and the measured distribution of black body radiation has led to numerous attempts to provide a theoretical basis for this formula. Of these the one receiving widest acceptance is that of Einstein,[1] who obtains the Planck formula by assuming merely that the chance of emission from an excited atom is increased by the presence of other light, of the same frequency as that which the atom emits, and that the increase is proportional to the concentration of that light.

I shall now approach this problem in a way which seems to me even more fundamental. The distribution of radiant energy in thermal equilibrium may be calculated as soon as we know the entropy of radiation as a function of concentration and of frequency. The notion of entropy may in turn

be referred to the still simpler notion of elementary probability. Without any premeditated attempt to obtain a particular result, I find that very simple assumptions regarding the probability lead directly to an expression for the entropy of radiation which is identical with the one obtained by Planck.

Adhering as closely as possible to the notation of Planck, the total radiant energy in a hollow (Hohlraum) will be designated by U, the energy density by $u = U/V$. The rate of change of energy density with frequency as we proceed through the spectrum will be called u_ν. Hence the density of energy lying between any two given frequencies is

$$\int_{\nu_1}^{\nu_2} u_\nu d\nu.$$

So also the total entropy, the entropy density, and the rate of change of entropy density with the frequency will be denoted by S, s, and s_ν. Planck makes the assumption that if the whole spectrum is divided into slices, each *slice* being that part of the radiation lying between two given frequencies, then the entropy of each slice, finite or infinitesimal, is the same as if that slice were alone in the hollow, all the remaining radiation being removed. I shall not go quite so far as this, but shall make a less sweeping assumption, the true significance of which will become apparent as we proceed. *Assumption I.* By taking successive intervals of frequency the whole radiation may be divided into slices so small that such quantities as u_ν and s_ν may be taken as constant throughout each slice, and yet so large that the entropy of each slice may be regarded as unchanged if all other radiation is removed.

As was first done by Joffé,[2] we may divide u_ν by $h\nu$ to obtain a new function $n_\nu = u_\nu/h\nu$, and from this obtain the related quantities, n and N. To anyone who has accepted the existence of light quanta or photons[3] these new quantities N, n and n_ν will represent, respectively, the number of these particles, their concentration and the rate of change of this concentration with the frequency. Even those who have not accepted the existence of these particles as proved will permit us to use these terms in order to see the consequences that may follow from the employment of a method analogous to that employed in the study of the entropy of gases.

The problem that we have to solve is this: If we have a certain amount of radiation in a given enclosure, what is the change in entropy when this radiation is allowed to escape freely into a larger enclosure? This problem may be put in another way: Suppose that we have N particles in a large enclosure, what is the chance of finding at any instant that all of these particles are contained in a given portion of this enclosure?

If V' is the volume of the large enclosure, and V the smaller volume of the selected portion, and if there is only one particle present, the chance

that it will happen to be in the small volume is V/V'. Now if instead of one particle we have N particles, the simplest assumption is that the chance that any one particle will be in the volume V is independent of the number of other particles already there, in which case the chance that all the particles will be in the smaller volume is $(V/V')^N$, and applying the principle of Boltzmann, the difference in entropy between N particles in the volume V, and N particles in the volume V' is given by the expression

$$S - S' = k \ln (V/V')^N = kN \ln (V/V'). \qquad (1)$$

This is the same as the expression for the change of entropy in the free expansion of a perfect gas. This equation for the dependence of the entropy of radiation upon the volume leads by known methods, which I shall not repeat here, to a formula for the distribution of radiant energy of the type of Wien's, which, as we know from experiment, is nearly but not quite in accordance with the truth. We also know that the corresponding expression for the entropy of a gas is not correct for any actual but only for an ideal gas.

We are, therefore, led to abandon the simplest assumption that the probabilities for the several particles are independent, and to try the following which is obviously the next simplest assumption. *Assumption II.* The chance of any one particle being in a selected volume V is a linear function of the number of particles already in that volume. Assumptions I and II suffice for a complete solution of the problem.

If then N particles are supposed to be enclosed in the fixed volume V', the chance that they will all be in a selected portion of this volume V is $W = W_1 W_2 W_3 \ldots W_N$, and the individual probability for any one particle, numbered $m + 1$, will be given by the equation,

$$W_{m+1} = \frac{V}{V'} + \frac{ma}{V'} \qquad (2)$$

where V' and a are constants. If we had been dealing with the molecules of a gas which have no effect on one another, except that each has a finite volume, we should find this same equation, in which the constant a would have a negative value. In our present case we shall regard a as completely undetermined as to sign and magnitude.

The probability that all N particles will be in the volume V is, therefore,

$$W = \frac{V}{V'} \frac{V+a}{V'} \cdots \frac{V+(N-1)a}{V'} = \left(\frac{a}{V'}\right)^N \left[\frac{V}{a}\left(\frac{V}{a}+1\right)\cdots\left(\frac{V}{a}+N-1\right)\right]. \qquad (3)$$

Our problem is not changed if, keeping V', V and N in the same ratios, we make them as large as we please. We may, therefore, make V/a very large and, without loss of generality, an integer, so that

$$W = \left(\frac{a}{V'}\right)^N \frac{\left(\frac{V}{a} + N - 1\right)!}{\left(\frac{V}{a} - 1\right)!}. \quad (4)$$

Now from the Boltzmann equation for the difference in entropy of N molecules in V and N molecules in V',

$$S - S' = k \ln W, \quad (5)$$

and we may employ the Stirling formula, neglecting those terms which become rapidly negligible when V/a and N become very large, in the form

$$\ln x! = x \ln x - x \quad (6)$$

and also neglecting unity in comparison with V/a and N, we find

$$S - S' = kN \ln \frac{a}{V'} + k\left(\frac{V}{a} + N\right) \ln \left(\frac{V}{a} + N\right) - k\frac{V}{a} \ln \frac{V}{a} - kN. \quad (7)$$

The quantity a has less effect upon the expression of entropy, the larger the volume is; as we may consider that V' was originally taken so large compared with V that the entropy S' is in the region of validity of equation (1). Hence, we may write

$$S' = kN \ln \frac{V'}{Nb}, \quad (8)$$

where Nb is a constant analogous to an integration constant, and may for the present be regarded as undetermined.

Combining (7) and (8), and performing a mere algebraic rearrangement, we find

$$S = kN \ln \frac{a}{be} + \frac{kV}{a}\left[\left(1 + \frac{Na}{V}\right) \ln \left(1 + \frac{Na}{V}\right) - \frac{Na}{V} \ln \frac{Na}{V}\right], \quad (9)$$

where e is the base of the natural logarithms. Dividing by the volume, we obtain the expression for the entropy density as a function of the concentration, $n = N/V$,

$$s = kn \ln \frac{a}{be} + \frac{k}{a}\left[(1 + na) \ln (1 + na) - na \ln na\right]. \quad (10)$$

So far it has not been necessary to be particular as to the kind of radiation present in the hollow or as to whether the correction factor a is the same when the particles are all of nearly the same frequency or of quite different frequencies. But now, in order to inspect more closely the meaning of this correction factor we may consider that we have present in the hollow only one of our slices of radiation, lying between the fre-

quencies ν and $\nu + \Delta\nu$. Then by Assumption I, $s_\nu = s/\Delta\nu$ and $n_\nu = n/\Delta\nu$, and equation (10) becomes

$$s_\nu = kn_\nu \ln \frac{a}{be} + \frac{k}{a\,\Delta\nu}\left[(1 + n_\nu a\,\Delta\nu)\ln(1 + n_\nu a\,\Delta\nu) - n_\nu a\,\Delta\nu \ln n_\nu a\,\Delta\nu\right]. \quad (11)$$

We may now introduce the thermodynamic principle embodied in Wien's displacement law. From any of the familiar expressions for that law we may readily obtain the equation,

$$s_\nu = n_\nu f\left(\frac{n_\nu}{\nu^2}\right) \text{ or } s_\nu = \nu^2 F\left(\frac{n_\nu}{\nu^2}\right). \quad (12)$$

Hence by inspection of any one of the terms of (11) we see that $n_\nu a\,\Delta\nu$ is a function only of n_ν/ν^2, and, therefore, a is proportional to $1/\nu^2\Delta\nu$. The factor of proportionality is now an absolute constant independent of any of the variables which determine the state of the system. It will, however, depend upon the units in which these variables are expressed, but this dependence may also be eliminated if we note that a has the dimensions of (length),[3] while $1/\nu^2\Delta\nu$ has the dimensions of (time).[3] Introducing c, the velocity of light, we have

$$a = \frac{\alpha c^3}{\nu^2\Delta\nu}, \quad (13)$$

where α is now independent of the variables and the units. It is, therefore, a pure number. We also see by inspection, comparing equations (11) and (12), that b must depend upon the variables in the same way as a, and, therefore, if β is another pure number,

$$b = \frac{\beta c^3}{\nu^2\Delta\nu}. \quad (14)$$

Equation (11) now becomes

$$s_\nu = kn_\nu \ln \frac{\alpha}{\beta e} + \frac{k\nu^2}{\alpha c^3}\left[\left(1 + \frac{n_\nu \alpha c^3}{\nu^2}\right)\ln\left(1 + \frac{n_\nu \alpha c^3}{\nu^2}\right) - \frac{n_\nu \alpha c^3}{\nu^2}\ln\frac{n_\nu \alpha c^3}{\nu^2}\right]. \quad (15)$$

While we have thus acquired information regarding the quantity a by purely mathematical methods, it will be interesting to translate this information into physical terms. Assumption I states that the entropy of one slice of radiation is independent of the presence of other slices. Therefore, any statement regarding the probability that the particles of a given slice be in a certain volume cannot be affected by the presence or absence of particles belonging to other slices. On the other hand, it can-

not be assumed that the particles are affected only by other particles of precisely the same frequency, for presumably no two particles have frequencies which are exactly the same. We are thus forced to regard the mutual effect of two particles as a sort of resonance phenomenon, diminishing with great rapidity as the difference in frequency between the two particles increases. (It seems probable that this mutual effect would almost entirely disappear with a difference of frequency as great as the width of a fine spectral line.)

Since we are about to find that a is positive, we might be tempted to regard this mutual effect of two particles as analogous to attraction. Indeed, I was formerly tempted to assume an average diminution in the velocity of a particle in the presence of other like particles; or, in other words, that the refractive index of a space containing radiation is slightly higher than unity. It is, perhaps, possible that such a phenomenon may be found if we ever are able to study radiation of enormous density, such as exists in the interior of the stars. But I am now convinced that it is entirely negligible in the region of our present measurements. If there were such a change in the velocity of radiation in the presence of other radiation, most of our thermodynamics of radiation, including the Stefan-Boltzmann law and Wien's displacement law, would fall to the ground.

I feel that the mutual effect of one particle upon the probability of the presence of another particle is a more subtle thing than can be accounted for by the mere assumption of forces acting between the particles *after they have started* from emitting atom to absorbing atom. But this is a matter which I cannot discuss further here, but must postpone until I have developed further my new theory of light.[4]

Returning to equation 15, we may substitute $u_\nu = h\nu n_\nu$ and find

$$s_\nu = \frac{ku_\nu}{h\nu} \ln \frac{\alpha}{\beta e} + \frac{k\nu^2}{\alpha c^3}\left[\left(1 + \frac{\alpha c^3 u_\nu}{h\nu^3}\right) \ln \left(1 + \frac{\alpha c^3 u_\nu}{h\nu^3}\right) - \frac{\alpha c^3 u_\nu}{h\nu^3} \ln \frac{\alpha c^3 u_\nu}{h\nu^3}\right]. \quad (16)$$

This is the equation obtained by Planck if we assume that

$$\alpha = \frac{1}{8\pi}. \quad (17)$$

This assumption completes the definition of the interesting quantity a. To obtain precisely the Planck form it would be necessary to assume also that $\alpha = \beta e$. This assumption, however, is unimportant and probably unnecessary. It affects only the additive term in the entropy equation, and, as we shall see in a further development of the subject, there is no inconvenience in leaving β undetermined for the present.

It is interesting to note that not a but $a\,\Delta\nu$ is specific, or, in other words independent of the width of the slice considered. If we now substitute the new value of α and write

$$\frac{a\Delta\nu}{c} = \frac{c^2}{8\pi\nu^2}, \tag{18}$$

we have a quantity which is of the dimensions of an area. Now there have been several attempts to calculate what, for brevity, has been called the cross-section of the quantum,[5] and the one by Dr. Smith and myself, based on a classical formula of Lord Rayleigh, gave just this expression $c^2/8\pi\nu^2$. I do not wish to stress this point here, but merely allude to it in passing.

Finally, it may be pointed out that, owing to the peculiar phrasing of Assumption I, we have obtained the Planck equation for entropy not as an exact formula but as an approximation; but certainly, admitting the general validity of the methods here employed, it follows that if equation 1 represents the first approximation to the law of entropy, equation (16) represents the obvious and necessary second approximation.

[1] Einstein, *Physik. Z.*, **18,** 121 (1917).

[2] Joffé, *Ann. Physik*, **36,** 534 (1911).

[3] Lewis, *Nature*, **118,** 874 (1926).

[4] Lewis, these PROCEEDINGS, **12,** 22 (1926).

[5] Ornstein and Burger, *Zs. Physik*, **20,** 345 (1924); Lewis and Smith, *J. Amer. Chem. Soc.*, **47,** 1508 (1925).

---

*LINEAR ELEMENTS OF THE ELECTROMAGNETIC PINHOLE GRAPHS**

BY CARL BARUS

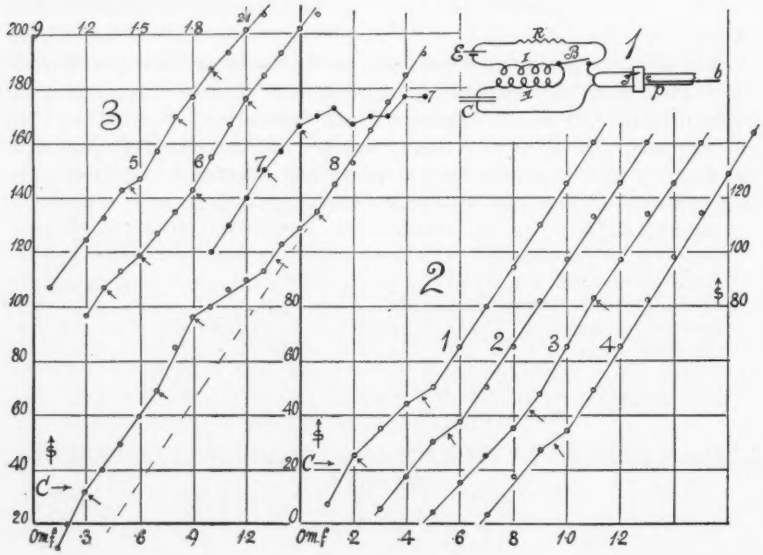
DEPARTMENT OF PHYSICS, BROWN UNIVERSITY

Communicated April 4, 1927

Following the suggestion at the end of the last paper[1] a new adjustment of apparatus was chosen, in which the condenser, C (Fig. 1), is charged directly on open circuit. Here B is the spring break (conveniently kept in resonance with the lighting circuit in the key of B), E and R electromotive force (2 cells) and resistance, T, p, telephone and organ pipe, I, II, primary and secondary of the transformer. When B is open, C is charged by E and discharged on closing. The coils I and II were eventually to be removed. The pinhole probe bc has its point about 1 cm. from the bottom of the pipe p.

Figure 2 shows the s, C graphs for the intervals 0 to 1.1 m.f. These

are successive measurements, each graph (1, 2, 3, 4) being in turn moved 0.2 m.f. to the right for clearness. The pipe was carefully tuned for the largest s available, in all cases. The results are a set of data strikingly linear and parallel in their main features, except for the occurrence, almost capriciously, of the breaks indicated by the arrows. As the spring interruptor was kept in tune (beats) the breaks in question are results of an accidental intonation of another harmonic, as soon as the conditions for it are at hand, or the instability of the original harmonic is excessive. This happens for capacities below 0.5 m.f., while 0.4 m.f. usually suffice for the breakdown.



It seemed desirable to determine how far this behavior would be prolonged and in figure 3, the graphs 5, 6, 7 are worked out between $C = 0.9$ and 2.2 m.f., and in number 8, between 0 and 2.2 m.f. The tendency to linear and parallel variation persists; but the breaks occur far more capriciously, as one would expect for these regions of low pitch. No. 7 after $C = 1.5$ m.f., breaks almost to the horizontal. No. 8, where the whole interval is tried out, is interesting as the slope of the line at the lower and at the top end are about the same (see dotted line). One also notes that after these breaks the curve does not recover, but the s remains continuously below the prolongation of the lower part of the curve. This would also be expected since the fringe-displacement, s, measures the difference of level of the mercury surfaces of the U-gauge and hence the wave

energy potentialized at the pinhole. Each new increment is added to the stored energy or may also be withdrawn from it.

The results for cases in which the primary or secondary, or both primary and secondary are removed, the pipe being tuned for each case, separately, were unsatisfactory. With the primary only in circuit (secondary removed) a distinctly higher pitch was heard, though the pipe depth is about the same. In this case and when both *I* and *II* are cut out, the graphs have definite crests and are, as a whole, more curvilinear than the preceding. With *II* only in place, the largest fringe displacements, *s*, obtainable are too small to be of service. In fact, taken together, these graphs are, throughout, small in their *s* values, as compared with figures 2 and 3 with both *I* and *II* in place; and the latter, in turn of 4–5 times less sensitivity than the earlier graphs (loc. cit.) for a completely separated primary and secondary. The marked tendency to preserve linear progress in the present cases has, however, been put in evidence.

* Advance note from a Report to the Carnegie Institution of Washington, D. C.

[1] These PROCEEDINGS, **13**, 1927, p. 87.

THERMIONIC EMISSION AND THE "UNIVERSAL CONSTANT" A

BY EDWIN H. HALL

JEFFERSON PHYSICAL LABORATORY, HARVARD UNIVERSITY

Communicated March 30, 1927

This paper continues my endeavor to show that certain ideas embodied in the dual theory of electric conduction within metals are capable of illuminating the field of electronic emission.

It is probable that nearly every one who has undertaken to discuss the behavior and functions of "free" electrons within metals has begun by assuming these electrons to share the energy of thermal agitation. This assumption is suggested by the general kinetic theory of the solid state, and is supported by the behavior of electrons that have been emitted by hot metals. It is embodied in the ordinary formulas, differing from each other only by a constant factor, for electric conduction by means of free electrons within metals. Only very powerful evidence or argument should be accepted as upsetting this general pre-supposition.

Nevertheless, workers in thermionic emission have seemed disposed, too hastily I think, to admit this disturbing change of conception. Accordingly, I shall here present and comment upon the various considerations that, so far as I know, have been most influential in forming opinion on the question at issue:

(*a*) Richardson, on p. 34 of his "Emission of Electricity" (1916), says that "the optical properties of metals lead us to conclude that the number of free electrons present in them is quite large, and if this large number of electrons possessed the kinetic energy which the classical dynamics endows them with, the specific heats of metals would be very much larger than those actually observed." This is, I believe, substantially equivalent to saying that we cannot explain metallic electric conduction by means of free electrons only without getting into difficulties over specific heats.

But the dual theory of conduction attributes a relatively small share of this operation to the free electrons. Moreover, the general conception as to the effectiveness of a single free electron in conduction has, I think, changed since the passage quoted from Richardson was written. I believe, and I am not alone in believing, that the conclusions to be drawn from the optical properties of metals are by no means unmistakable.

(*b*) Experiments like those of Davisson and Germer[1] indicate that the amount of energy needed to maintain a given thermionic emission current is greater than it would be if all the emitted electrons entered the emitting filament as free electrons sharing the energy of thermal agitation.

But I have shown[2] that, according to the dual theory of conduction, most of the emitted electrons *enter* the emitting filament in the "associated" conductive state, without thermal energy. If this view of the matter is taken, there is no need of assuming that the free electrons also within the metal lack energy of thermal agitation. I believe that the same consideration will explain the evidence[3] which has inclined Millikan to the view that most of the conductive electrons do not have thermal energy.

(*c*) Finally we have the measure of success, whatever it may prove to be, attained by Dushman and others in calculating the value of the constant A, in Richardson's emission equation,

$$I = AT^2\epsilon^{-\frac{b_0}{T}}, \qquad (1)$$

by means of Nernst's "third law" and the quantum theory.

But Bridgman[4] has subjected Dushman's argument to a thorough criticism, and, though he is disposed to admit that the calculated value of A may prove to be approximately correct in the case of all pure metals, he maintains that not all of Dushman's conclusions are justified.

In particular, Bridgman argues that Dushman's success in the case of tungsten, where he treats the "surface heat" as zero, is no proof that this heat is really zero. His conclusion is given as follows: "It is therefore not necessary that the surface heat vanish in order that the emission formula hold [with Dushman's value of the 'universal constant' A]; numerically it may be large or small provided only that it be proportional to temperature."

Bridgman defines "surface heat" as "the heat which must be absorbed by the system to maintain it isothermal when one unit of electricity is added to the surface." Now, according to the dual theory of conduction, as I showed[5] some years ago, this surface heat is due to the ionization which occurs at the surface, in accordance with the mass law of chemical equilibrium, when a positive charge (taking away of free electrons) is added there. Accordingly the surface heat at any temperature should be proportional to what I have called the heat of ionization within the metal and have indicated, per electron freed, by the symbol λ' in the equation

$$\lambda' = \lambda_c' + s\,k\,T, \tag{2}$$

given as equation (3′) in my "Summary"[6] printed six years ago. I have thus far taken λ_c' and s in this equation as constants, though it is possible, as will be shown later, that s should be regarded as a slowly changing function of temperature. The value of λ_c' given for tungsten in the paper referred to is $40k$, while that of s is 11.4. Accordingly, if we take T as 2000°, a reasonable temperature for emission experiments on tungsten, we find that the second term in the value of λ_c', which term is proportional to T, is 570 times the value of the constant λ_c'. I attach no great importance to the particular values of λ_c' and s here given for tungsten, as they were derived from somewhat conjectural data and should be regarded as tentative; but apparently equation (2), which I have used as one of the fundamental formulas of the dual-conduction theory, may, for approximate purposes at least, be regarded as indicating a "surface heat" proportional to the temperature and so, according to Bridgman's argument, held consistent with the practical correctness of equation (1), with Dushman's value of A, in the case of tungsten.

There is one other particular in which recent discussion seems favorable to one of my cherished notions. Equation (1) of the "Summary" already referred to is

$$n = zT^q, \tag{3}$$

where n is the number of free electrons per unit volume within a metal, while z and q are constants characteristic of the metal. Values of q as given in the "Summary" for eighteen metals, including two alloys, range from 1.2 for iron and bismuth to 1.6 for nickel, palladium and platinum, the average value being 1.47. Now the tendency in current treatment of emission is to establish Richardson's equation (1) of this paper, in preference to his older formula containing $T^{0.5}$ instead of T^2. But Richardson, on p. 34 of his "Emission of Electricity" (1916), writes that, "if n_1 [meaning the same as my n] is proportional to $T^{1.5}$ and ϕ is independent of T,

$$I = A_2 T^2 \epsilon^{-\frac{\phi_0}{kT}},\text{"} \tag{4}$$

provided we keep to the "classical kinetic theory" concerning conditions within the metal. That is, Richardson in the passage referred to recognized that, if the classical theory held, as he had apparently assumed at first, the acceptance of equation (4) implied the relation

$$n = zT^{1.5}, \tag{5}$$

for the number of free electrons per unit volume within a metal, z being a constant. Comparing this with my equation (3) above, and remembering that my mean value of q is very near 1.5, we find a degree of coincidence that is rather impressive, especially in view of the fact that thermionic literature in general gives little prominence to the idea that n increases with rise of temperature.

What precedes is enough, I trust, to show that the dual-conduction theory, including the classical kinetic conception as to the state of free electrons within a metal, should be reckoned with in future discussions of electronic emission. Furthermore, it should not be forgotten that this theory gives very definite and simple explanations of the Volta effect and of the Peltier effect, including that phase of it which Bridgman was surprised to find within a single metal crystal. In addition it makes possible a clear definition of "electron affinity," otherwise a rather vague term.

But I can well go farther than this. Equation (1) of this paper is recognized by Richardson and also by Bridgman as probably only an approximation to the true formula of thermionic emission. I shall, therefore, keeping to the classical kinetic conception of the free electrons within a metal, and following in general that particular method of Richardson which begins on p. 33 of his "Emission of Electricity" (1916), undertake to derive and discuss a form of emission equation that will embody those concepts of the dual-conduction theory which find application here.

If we assume a condition of equilibrium between the free electrons within a metal and the free-electron atmosphere outside the metal, we have a perfectly simple case for the application of Boltzmann's distribution law. That is, letting n be the number of electrons per unit volume in the surface layer of the metal and n' the number in unit volume of space just outside the surface, and defining "electron affinity" in this case as the amount of work, w, done against resisting forces in the passage of one electron from the free state within to the free state just outside the metal, we have

$$\frac{n'}{n} = \epsilon^{-\frac{w}{kT}}. \tag{6}$$

For n I shall substitute zT^q, according to equation (3) above. As to w, I have in these PROCEEDINGS for August, 1926, given my reasons for regarding it as a constant, which I will here write as b_0k, minus $(s-2.5)kT$,

where s is the same as in equation (2) of the present paper. That is, putting a for $(s-2.5)$, we have

$$w = k(b_0 - aT). \tag{7}$$

Thus from (6) we get, by means of (3) and (7),

$$n' = zT^q\epsilon^{-\frac{(b_0-aT)}{T}} = \epsilon^a zT^q\epsilon^{-\frac{b_0}{T}}. \tag{8}$$

According to Richardson's admittedly venturesome but generally accepted assumption, there is no reflection of electrons from the outer surface of the metal, so that N', the total number of electrons entering unit area of the surface per second from without, is the same as the number reaching this area per second from without. This number is, according to the kinetic theory of gases, proportional to n' and to the square root of the absolute temperature; that is,

$$N' = Fn'T^{0.5}, \tag{9}$$

where F is a constant, $(k \div 2\pi m)^{0.5}$, m being the mass of an electron. The numerical value of F is 1.56×10^5.

Now in the case of equilibrium this inflow must be offset by an equal outflow. It is, furthermore, assumed by Richardson and accepted by others that, when equilibrium is upset by applying an external potential-gradient sufficient to prevent any outside electrons from returning to the metal, the outflow of electrons from the surface is the same as in the case of equilibrium conditions. Argument in favor of this assumption is found in the fact that moderate increase of the potential-gradient beyond a certain value does not increase the observed outflow. Accordingly, the saturation current outflow, I, measured for unit surface area, is equal to e times the N' of equation (9). That is,

$$I = Fen'T^{0.5}. \tag{10}$$

From (10) and (8) we get

$$I = Fe\epsilon^a zT^{(q+0.5)}\epsilon^{-\frac{b_0}{T}}. \tag{11}$$

Putting A' for the product $Fe\epsilon^a z$, we get from (11)

$$I = A'T^{(q+0.5)}\epsilon^{-\frac{b_0}{T}}. \tag{12}$$

Equation (12) differs from equation (1) in having $(q + 0.5)$ instead of 2 as the exponent of the T factor and in having instead of the constant factor A the factor A', which may prove to be not altogether constant. I shall presently discuss these differences at some length, but for the moment I wish to comment on the constant b_0. This may be numerically the same as Richardson's b_0, each being defined as $(1 \div k)$ times the value

of w for $T = 0$. But Richardson (loc. cit.) in arriving at equation (1) assumes w, his ϕ, to be independent of temperature, whereas I take its value as expressed by equation (7).

It has become the fashion to discuss the claim of A in equation (1) to be regarded as a "universal constant," the same for all pure metals. My contribution to this discussion must involve a somewhat roundabout argument. Going back to equation (11), which immediately preceded the introduction of my A', I can, using equation (3), write

$$I = Fe\epsilon^a zT^q \cdot T^{0.5}\epsilon^{-\frac{b_0}{T}} = Fe\epsilon^a nT^{0.5}\epsilon^{-\frac{b_0}{T}}, \qquad (13)$$

where Fe is a constant.

I shall deal with the question whether the factor $Fe\epsilon^a zT^q$, or its equivalent $Fe\epsilon^a n$, is likely to be the same in different metals at the temperature T. If I took q to have the same value, 1.5, for all metals, I should in this inquiry be asking whether $Fe\epsilon^a z$, which is the A' of equation (12), is a universal constant.

The definition which I have given to a is such that I can rewrite equation (2) as

$$\lambda' = \lambda_c' + (a + 2.5)kT. \qquad (14)$$

In this equation, $2.5kT$ is the heat energy which one electron, after the process of ionization, possesses as a gas particle within the metal. If we denote by γ the amount of kinetic energy which this electron possessed before the ionization and by i the amount of work done against resisting attractions or repulsions in freeing the electron from the atom of which it formed a part, we shall have

$$i = \gamma + \lambda_c' + akT. \qquad (15)$$

The quantity i, if it needs a name, may be called the work-function of ionization within the metal. It is unnecessary for our present purpose to attempt any evaluation of γ. That part of it which represents energy of thermal agitation is negligibly small, and accordingly γ as a whole may be treated as a constant. It is to be noted in passing that γ may prove to be the larger part, at ordinary temperatures, of i in equation (15).

It is one of the tenets of the dual-conduction theory that, when any two metals α and β are joined at the same temperature, the total energy, including the energy γ, of an "associated" electron, an electron capable of taking part in what I call associated-electron conduction, is the same in an atom of α as in an atom of β. This tenet[7] was not first suggested by but is strongly supported by the observed fact that, when α and β are joined together, the photo-electric work function is the same for both.[8]

Accordingly, the total amount of energy required from without to take an associated electron from an atom of β to the free state in β, thence to

the free state in α and thence to the associated state in α, is zero. Let the work required for the second step of this process be called $w_{\beta\alpha}$. Then we have

$$(\lambda_c' + (a + 2.5)kT)_\beta + w_{\beta\alpha} - (\lambda_c' + (a + 2.5)kT)_\alpha = 0, \quad (16)$$

or

$$w_{\beta\alpha} = (\lambda_c')_\alpha - (\lambda_c')_\beta + (a_\alpha - a_\beta)kT. \quad (17)$$

Now, as a rule, according to the values given in my "Summary," already referred to more than once, λ_c' is decidedly smaller than akT for any metal, even at room temperature, though not to be ignored. At thermionic emission temperatures, as I have already shown in connection with equation (2) in this paper, the value of λ_c' may in the case of some metals be negligibly small compared with skT and even compared with akT. If we can in the case of α and β at high temperatures neglect the λ_c' terms in the value of $w_{\beta\alpha}$, we have from (17)

$$w_{\beta\alpha} = (a_\alpha - a_\beta)kT. \quad (18)$$

But evidently, n_α being the number of free electrons per unit volume in α and n_β the corresponding number in β, we have by the Boltzmann law

$$\frac{n_\alpha}{n_\beta} = \epsilon^{-\frac{w_{\beta\alpha}}{kT}}, \quad (19)$$

and so, from (18)

$$\frac{n_\alpha}{n_\beta} = \epsilon^{-(a_\alpha - a_\beta)}, \quad (20)$$

or

$$(\epsilon^a n)_\alpha = (\epsilon^a n)_\beta. \quad (21)$$

Accordingly, since the Fe of equation (13) is a universal constant, the quantity $Fe\epsilon^a n$, or $Fe\epsilon^a zT^q$, of the same equation is, within the limits of accuracy of the preceding argument, the same for metal α as for metal β.

It is to be noted that, if a is zero and if q is 1.5, equation (13) reduces to the form

$$I = FezT^2\epsilon^{-\frac{b_0}{T}}, \quad (22)$$

which is equivalent to the Richardson equation that I have written as (1). Richardson does, in fact, in that part of his "Emission" to which I have referred, take a as always zero and q as always 1.5 in getting this equation. Accordingly, his A is my Fez of equation (22), and if A is to be taken as a universal constant, Fe being undoubtedly such, z must also be a universal constant, and so $zT^{0.5}$, the number of free electrons per unit volume,

according to Richardson's assumptions here, must be the same for all metals at the same temperatures. This seems an improbable conclusion. We can avoid it in any one of three ways. We can give up the simple classical kinetic conception as to the state of the free electrons within the metal, in spite of the great advantages it offers in several respects, or we can assume either that a is not always zero or that q is not always 1.5. I have made both of these saving assumptions.

In neither case was I influenced originally by any consideration of electronic emission. My assumption regarding q, as introduced in my adoption of equation (3), was made in an endeavor to account for the known facts of thermo-electric action. I brought in a, by making the value of s in equation (2) greater than 2.5, for the purpose of enabling my theory of dual electric conduction to account for the small temperature coefficient of thermal conductivity as compared with the temperature coefficient of electric conductivity. If I should be compelled to give up the attempt to explain the experimentally known relations of these two conductivities and to adopt the conception that thermal conduction in metals is carried on wholly or mainly by means of elastic vibrations, I might drop the assumption that a is not zero. But I feel very sure that in any case the conception of q as not always 1.5 would remain, so long as I might try to understand thermo-electric action.

With given values of I and b_0 the effect of introducing a, which is always positive, is to reduce in equation (13) the factor zT^q, which represents the number of free electrons per unit volume within the metal. This is as it should be; for the greater a is, with a fixed value of b_0, the less is the work, see equation (7), required to remove a free electron from inside to outside the metal, and evidently the less this work the smaller the number of interior free electrons required to maintain a given outflow I.

The particular values of a and q which I have published in my "Summary" are entitled to no great confidence, even if my dual-conduction theory is accepted as sound, for they were derived, as I have explained,[9] from imperfect data taken for the temperature range from 0°C. to 100°C. and by a method lacking in strict accuracy. Nevertheless, I shall undertake to apply them in one or two cases, largely for the purpose of finding what values of n are likely to be indicated by the logic of this paper.

The constant Fe, which must appear in every case, has the numerical value

$$1.56 \times 10^5 \times 1.59 \times 10^{-20} = 2.48 \times 10^{-15}.$$

Tungsten.—I have $q = 1.48$ and $a = 11.4\text{–}2.5 = 8.9$. From data given by Dushman[10] and his co-workers it appears that, in the case of tungsten, if I is measured in absolute electromagnetic units, the b_0 of equation (1) is about 52,640 and the A is about 6.02. In order, with the same value of

b_0, to make my equation (13) conform to the experimental data, I must have

$$A'T^q = AT^{1.5} = 6.02T^{1.5},$$

whence, if we, for example, take $T = 2000$,

$$A' = 6.02 \times 2000^{0.02} = 7.01.$$

Then, from the definition of A',

$$z = 7.01 \div Fe\epsilon^a = 7.01 \div (2.48 \times 10^{-15} \times 2.72^{8.9}) = 3.86 \times 10^{11},$$

and, for $T = 2000$,

$$n = zT^{1.48} = 2.97 \times 10^{16}.$$

Molybdenum.—I have here $q = 1.3$, $a = 8-2.5=5.5$. From Dushman's paper I find that the universal constant A is not quite so constant as one might expect from its name. Its $\log_{10}$ is given as 1.776 in the case of tungsten and 2.268 for the case of molybdenum, when I is measured in amperes. When I is measured in absolute units, we have for A in the case of molybdenum the value of 18.5, about three times as great as for tungsten.

Accordingly, for molybdenum we get, taking T as 2000,

$$A' = 18.5 \times 2000^{(1.5-1.3)} = 84.8.$$

Then

$$z = 84.8 \div (2.48 \times 10^{-15} \times 2.72^{5.5}) = 1.40 \times 10^{14}.$$

And, for $T = 2000$,

$$n = zT^{1.3} = 1.40 \times 10^{14} \times 2000^{1.3} = 2.74 \times 10^{18}.$$

I have no estimates of q and a for tantalum, the third metal dealt with by Dushman.

The value of n at 2000° which I have arrived at by the method above shown is about ninety times as great in the case of molybdenum as in the case of tungsten. This ratio seems improbably large, in view of the fact that the electric conductivities of the two metals are not very different; but when due consideration is given to the possibilities of error in the valuations of a, q and A, the ratio is not large enough to discredit seriously the argument that led to it. I shall have something more to say on this matter at the end of this paper.

On the other hand, even the value of n found for molybdenum at 2000° is surprisingly small, about 1 in 20,000 as compared with the number of the atoms. From one point of view this outcome is gratifying, as, if accepted, it banishes all concern about the influence of the free electrons on the specific heat of metals. But it raises at once the question whether so small a number of electrons, and the corresponding number of ions

can reasonably be supposed to provide the observed value of electric conductivity.

This question may well be regarded as an open one at present. The answer will depend upon the admissible length of that very dubious quantity the "mean free path" of the free electrons within the metal and on the conception which may ultimately prevail as to the exact manner in which the ions function in facilitating the passage of electrons which are not free in the ordinary sense. I do not insist, nor have I ever insisted, on any particular mode as the one in which this influence must be exercised, though I have more than once made suggestions which seemed to me at the time to have some degree of plausibility regarding this matter. What I have emphasized and would still emphasize is the great importance of the ions as furnishing leeway, clearance space, terminal facilities, whatever the freedom indicated may best be called, for the movement of "associated" electrons, singly or in files, to progress in the direction of the imposed electromotive force, without maintaining a lock-step relation to each other all around the circuit.

I have once or twice suggested in the preceding pages that the a of my equations may not be strictly a constant. In fact, if equation (21) be written in the equivalent form

$$(\epsilon^a z T^q)_\alpha = (\epsilon^a z T^q)_\beta, \tag{23}$$

we see at once that, z being taken as a constant for each metal, a_α and a_β cannot both be taken as constant unless $q_\alpha = q_\beta$. It is quite probable that both a and q are variables in the case of every metal; for it seems unlikely that equations so simple as (14), with λ_c' and a constant, and (3), with z and q constant, should hold all the way from room temperatures up to $T = 2000°$.

I shall, however, for the present assume z_α, q_α, z_β and q_β to be constants through any range of temperature here dealt with, and accordingly I shall put upon a the responsibility of changing with temperature in such a way as to maintain equation (23). This requires the differential relation

$$d(a_\beta - a_\alpha) = (q_\alpha - q_\beta) d \log T, \tag{24}$$

whence, if we integrate from $T = 325°$, approximately the temperature for which the conclusions of my "Summary" apply, to $T = 2000°$, we have

$$\int_{325}^{2000} d(a_\beta - a_\alpha) = (q_\alpha - q_\beta) \int_{325}^{2000} d \log T = 1.8\,(q_\alpha - q_\beta). \tag{25}$$

Of course, this change might come partly in a_α and partly in a_β, α indicating tungsten, let us say, and β indicating molybdenum, in which case $(q_\alpha - q_\beta) = 0.18$. I shall, for simplicity, assume that it occurs wholly

in a_β, which accordingly increases from the value 5.5, used above in the case of molybdenum, to 5.8 at 2000°. If with this new value of a I repeat my calculations relating to molybdenum, I have

$$z = 84.8 \div (2.48 \times 10^{-15} \times 2.72^{5.8}) = 1.03 \times 10^{14},$$

and

$$n = zT^{1.3} = 1.03 \times 10^{14} \times 2000^{1.3} = 2.16 \times 10^{18}.$$

Thus the ratio of n for molybdenum to n for tungsten, at 2000°, is reduced from a value about 90 as found above, to a value about 65.

If, furthermore, I were to take for Dushman's A in the case of molybdenum the same value, 6.02, which I, on Dushman's authority, used in the case of tungsten, the ratio of the two n's in question would come out about 22, a value that still seems large.

I have sometimes criticized, as others have, the argument by which Richardson, beginning on page 28 of his "Emission of Electricity" (1916) and ending on page 33, has deduced his emission equation. I have believed, and I still believe, this argument to contain a fallacy, failure to distinguish properly between the work done against resisting forces in removing a free electron from a metal and the absorption of heat which attends this action; but I have not found it easy to decide how much harm had resulted from this fallacy in this particular connection. In the present paper, while differing from Richardson as to certain explicit assumptions, I have found no fault with his logic; for that argument of his which I have here been following is different from the one referred to above.

The fact is that Richardson has in the book mentioned given three derivations of his emission equation, the (1) of this paper. The first is the one criticized, the third, running from page 35 to page 39, follows the quantum theory and reaches a result not very different from that found by Dushman in a similar way. The second is contained in a single page, beginning on 33 and ending on 34. Unfortunately, I think, the author felt obliged to give it up, because of a misgiving as to assigning thermal energy to the free electrons within a metal. The purpose of the present paper has been to show that such a misgiving lacks justification, and to modify, without rejecting, Richardson's "classical kinetic theory" of thermionic emission.

I predict that the A of Richardson's emission equation, equation (1) of this paper, will not prove to be the same for different metals. I propose as a substitute for this equation (1) the equation numbered (12) in this paper, the A' therein having the meaning given in the text. I hope that, when sufficient data concerning the electric conductivity, the heat conductivity and the thermo-electric behavior of metals are available, the factor $A'T^q$, or $Fe\epsilon^a zT^q$, will be found approximately the same for metals in general.

[1] *Physic. Rev.*, **20** (1922), p. 300.
[2] *Proc. Nat. Acad. Sci.*, **13** (1927), pp. 43–46.
[3] *Physic. Rev.*, **27** (1926), pp. 51–67.
[4] *Ibid.*, **27** (1926), pp. 173–180.
[5] *Proc. Nat. Acad. Sci.*, **9** (1923), pp. 207–211.
[6] *Ibid.*, **7** (1921), pp. 98–107.
[7] Hall, "Conditions of Electric Equilibrium, Etc.," *Ibid.*, **11** (1925), pp. 111–116.
[8] Richardson and Compton, *Phil. Mag., London*, **24** (1912), p. 575; Hennings, *Physic. Rev., Ithaca*, **4** (1914), p. 228; Millikan, *Physic. Rev.*, **18** (1921), p. 236.
[9] Illustrations of the "Dual Theory of Metallic Conduction," *Physic. Rev.*, **28** (1926), pp. 392–417.
[10] *Ibid.*, **25** (1925), pp. 356 and 357.

LAGRANGIAN FUNCTIONS AND SCHRÖDINGER'S RULE

By H. Bateman

California Institute of Technology

Communicated March 23, 1927

In a recent paper Schrödinger[1] has extended a rule, used by writers in the theory of gravitation, for deriving a stress energy tensor from a Lagrangian function and has illustrated its application in the case of the tensor which he has associated with the system of equations proposed by Gordon.[2]

We shall now apply the rule to various Lagrangian functions to see if it is generally applicable. Let (a_1, a_2, a_3, a_4) be the components of a typical 4-vector on which the Lagrangian function depends and let a_{mn} denote the derivative of a_m with respect to the coördinate x_n. The rule then states that the component T_{mn} of an associated stress-energy tensor is given by

$$T_{mn} = \sum \left[\sum_{l=1}^{4} a_{ln} \frac{\partial L}{\partial a_{ln}} + \sum_{l=1}^{4} a_{ml} \frac{\partial L}{\partial a_{nl}} + a_m \frac{\partial L}{\partial a_n} \right] - \delta_{mn} L \qquad \ldots\ldots \qquad (1)$$

where

$$\delta_{mn} = 0 \quad m \neq n$$
$$= 1 \quad m = n$$

and the summation $\sum$ extends over all the four vectors of type a.

Let us now apply this rule to the Lagrangian function

$$L = \frac{1}{2}(E^2 - H^2) \qquad (2)$$

in which

$$\mathbf{E} = \mathbf{e} + \mathbf{e}^*, \quad \mathbf{h} = \mathbf{h} + \mathbf{h}^*$$
$$\mathbf{e} = -\frac{1}{c}\frac{\partial \mathbf{a}}{\partial t} - \nabla\phi, \quad \mathbf{h} = \text{curl } \mathbf{a} \tag{3}$$
$$\mathbf{e}^* = -\text{curl } \mathbf{b}, \quad \mathbf{h}^* = -\frac{1}{c}\frac{\partial \mathbf{b}}{\partial t} - \nabla\omega.$$

The variational principle

$$\delta \int\int\int\int L\, dxdydzdt = 0 \tag{4}$$

is found to give rise to the Maxwellian equations

$$\text{curl } \mathbf{H} = \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t}, \quad \text{div } \mathbf{E} = 0$$
$$\text{curl } \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{H}}{\partial t}, \quad \text{div } \mathbf{H} = 0 \tag{5}$$

but Schrödinger's rule, as it stands, does not give a symmetric stress-energy tensor. We, therefore, modify the rule by writing

$$S_{mn} = \frac{1}{2}(T_{mn} + T_{nm}) \qquad \ldots\ldots \tag{6}$$

and regarding S_{mn} as the components of the stress-energy tensor. It is now found that this tensor is the difference[3] of the tensors associated with the fields $(\mathbf{e}, \mathbf{h})$ and $(\mathbf{e}^*, \mathbf{h}^*)$, respectively, in other words

$$S_{mn} = s_{mn} - s^*_{mn}.$$

If $e^* = 0$, $h^* = 0$, we get the correct stress-energy tensor associated with the field (e, h), but if $e = 0$, $h = 0$, we get $-s^*_{mn}$. This anomaly may be avoided by taking as our Lagrangian function

$$L = \frac{1}{2}(E_0^2 - H_0^2)$$

where

$$\mathbf{E}_0 = \mathbf{e} + i\mathbf{e}^*, \qquad \mathbf{H}_0 = \mathbf{h} + i\mathbf{h}^*$$

but the use of a complex Lagrangian function seems undesirable.

An interesting system of equations may be obtained by using the Lagrangian function

$$L = \frac{1}{2}(E^2 - H^2) + \lambda(a.b - \phi\omega)$$

in which λ is an arbitrary constant. In this case both electricity and magnetism are present.[4]

Gordon's equations relate to the case in which electricity is distributed throughout space and its distribution and motion are governed by certain

equations which are quite hard to solve. The equations are interesting because the principles of the conservation of energy and momentum are consequences of the equations derived from the Lagrangian principle and these equations include two equations satisfied by conjugate complex functions ψ and $\bar{\psi}$ which closely resemble the wave equations used by de Broglie and Schrödinger.

When there is no field[5] these wave-equations are of type

$$\Box^2\psi \equiv \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} = k^2\psi$$

where k is a constant and the question arises whether these equations possess any solutions which are finite and continuous throughout all space.

The question is immediately answered in the affirmative by the consideration of solutions of type

$$\psi = e^{i(lx + my + nz - wt)}.$$

When $k = 0$, there are certainly other solutions, for if a is real and $\sigma^2 > 1$ the wave-function

$$\psi = \frac{1}{x^2 + y^2 + (z - iac)^2 - c^2(t - i\sigma a)^2}$$

$$= \frac{x^2 + y^2 + (z + iac)^2 - c^2(t + i\sigma a)^2}{[x^2 + y^2 + z^2 - c^2t^2 + a^2c^2(\sigma^2 - 1)]^2 + 4a^2c^2[z - \sigma ct]^2}$$

does not become infinite for any real values of x, y, z and t. The real part of ψ also possesses this property.

It is easy to write down other systems of field equations giving continuous distributions of electricity and electric currents for which the principles of the conservation of energy and momentum are satisfied identically. A very simple set of equations of this type is derived from the Lagrangian function

$$L = \frac{1}{2}(E^2 - H^2) + \lambda(A^2 - \Phi^2)$$

in which λ is a constant and

$$\mathbf{H} = \text{curl } \mathbf{A}, \quad \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} - \nabla\Phi.$$

The associated equations are

$$\text{curl } \mathbf{H} - \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} = \lambda\mathbf{A}, \quad \text{div } \mathbf{E} = \lambda\Phi$$

and these give

$$\mathbf{A} + \text{div}\,\frac{1}{c}\frac{\partial\Phi}{\partial t} = 0, \quad \Box^2\mathbf{A} + \lambda\mathbf{A} = 0, \quad \Box^2\Phi + \lambda\Phi = 0.$$

Another interesting set of equations is derived from the Lagrangian function

$$L = -\frac{1}{2}(H^2 - E^2) - \rho[\Phi - \frac{1}{c}(\mathbf{v}.\mathbf{A})] - \mathbf{p}.\mathbf{P} + \mathbf{q}.\mathbf{Q} + \frac{1}{2}\lambda^3[L^2 - \Lambda^2]$$

in which λ is a constant and

$$\mathbf{H} = \text{curl } \mathbf{A}, \quad \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} - \nabla\Phi$$

$$\mathbf{p} = \text{curl } \frac{\rho\mathbf{v}}{c}, \quad \mathbf{q} = -\frac{1}{c^2}\frac{\partial}{\partial t}(\rho\mathbf{v}) - \nabla\rho$$

$$\mathbf{P} = \text{curl } \mathbf{L}, \quad \mathbf{Q} = -\frac{1}{c}\frac{\partial \mathbf{L}}{\partial t} - \nabla\Lambda.$$

The associated equations are

$$\text{curl } \mathbf{P} - \frac{1}{c}\frac{\partial \mathbf{Q}}{\partial t} = \mathbf{A}, \quad \text{div } \mathbf{Q} = \Phi$$

$$\text{curl } \mathbf{H} - \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} = \frac{\rho\mathbf{v}}{c}, \quad \text{div } \mathbf{E} = \rho$$

$$\text{curl } \mathbf{p} - \frac{1}{c}\frac{\partial \mathbf{q}}{\partial t} = \lambda^3\mathbf{L}, \quad \text{div } \mathbf{q} = \lambda^3\Lambda$$

and these give

$$\text{div } \mathbf{A} + \frac{1}{c}\frac{\partial \Phi}{\partial t} = 0$$

$$\text{div } (\rho\mathbf{v}) + \frac{\partial \rho}{\partial t} = 0$$

$$\text{div } \mathbf{L} + \frac{1}{c}\frac{\partial \Lambda}{\partial t} = 0$$

$$\Box^2\mathbf{L} = -\mathbf{A}, \quad \Box^2\Lambda = -\Phi$$

$$\Box^2\mathbf{A} = -\frac{\rho\mathbf{v}}{c}, \quad \Box^2\Phi = -\rho$$

$$\Box^2\left(\frac{\rho\mathbf{v}}{c}\right) = -\lambda^3\mathbf{L}, \quad \Box^2\rho = -\lambda^3\Lambda.$$

An interesting feature of this system of equations is that when ρ and $\rho\mathbf{v}$ have been obtained by solving the equations

$$\Box^6\left(\frac{\rho\mathbf{v}}{c}\right) = -\lambda^3\frac{\rho\mathbf{v}}{c}, \quad \Box^6\rho = -\lambda^3\rho,$$

$$\text{div } (\rho\mathbf{v}) + \frac{\partial \rho}{\partial t} = 0$$

the Hertzian vectors $\mathbf{P}$, $\mathbf{Q}$ and the field vectors, E, H can be derived by differentiations and the potentials A and Φ are connected by the well-known relation

$$\operatorname{div} \mathbf{A} + \frac{1}{c}\frac{\partial \Phi}{\partial t} = 0.$$

In both of these cases the stress-energy tensor may be obtained by Schrödinger's rule.

[1] E. Schrödinger, *Ann. Phys.*, **82** (1927), p. 265. A special form of the rule is suggested immediately by Abraham's tensor. *Phys. Zeitschr.*, **13** (1912). Hilbert's rule, used in *Math. Ann.*, **92** (1924), p. 1, points to the existence of a more general rule.

[2] W. Gordon, *Zeitschr. Phys.*, **40** (1926), p. 117.

[3] The tensor thus vanishes completely when the potentials $\mathbf{a}$, ϕ, $\mathbf{b}$, ω are chosen so that $\mathbf{e} = \mathbf{e}^*$, $\mathbf{h} = \mathbf{h}^*$. It should be remarked that another tensor possessing this property may be derived from the Lagrangian function $L = \mathbf{E}\cdot\mathbf{H}$, which also gives rise to the field-equations (5).

[4] Another Lagrangian function which is consistent with the existence of both electricity, and magnetism is discussed by E. T. Whittaker, *Proc. Roy. Soc. London*, **113A** (1927), p. 496.

[5] The question whether this is the total field or only the external field may at present be left open. If there is to be no total field and consequently no electricity both ψ and $\bar{\psi}$ must be real.

TWO REMARKS ON THE WAVE-THEORY OF MECHANICS

BY F. D. MURNAGHAN AND K. F. HERZFELD

JOHNS HOPKINS UNIVERSITY

Communicated April 15, 1927

A. *On the Degree of Arbitrariness in the Wave-Equation.*—For a dynamical system with fixed constraints, moving in a stationary conservative force field, the action W has the form $W = -Et + S(q)$; here $E = T + V$ is the total energy, both kinetic and potential, of the system and the symbol q stands for the positional coördinates of the system. The function W of (q, t) satisfies the Hamilton equation $\frac{\partial W}{\partial t} + H\left(q, \frac{\partial W}{\partial q}\right) = 0$ where H is the Hamiltonian function whose numerical value is the constant E. If we wish to set up the partial differential equation of the family of moving surfaces $F(q, t) = 0$ which is characterized by the fact that on each member of the family W has a constant value (it being understood that the dynamical system is started off from any point but always with the same value of the energy constant E) we proceed as follows. F is to be a

function of W, $F = F(W)$, so that $\frac{\partial F}{\partial q} = F' \frac{\partial W}{\partial q}$, $\frac{\partial F}{\partial t} = F' \frac{\partial W}{\partial t} = -EF'$, where F' stands for the derivative of F with respect to W. Eliminating F' from these equations we have $\frac{\partial W}{\partial q} = -E \frac{\partial F}{\partial q} \Big/ \frac{\partial F}{\partial t}$ and on substituting these values in the Hamilton equation we find the equation satisfied by F; it is

$$H\left(q, -E\left(\frac{\partial F}{\partial q} \Big/ \frac{\partial F}{\partial t}\right)\right) = E.$$

On using the fact that $H = T + V$ where T is homogeneous of degree two in the momenta, which do not appear at all in V, this equation takes the familiar form

$$T\left(q, \frac{\partial F}{\partial q}\right) = \frac{(E-V)}{E^2}\left(\frac{\partial F}{\partial t}\right)^2.$$

The formal expression for the kinetic energy T is that which gives it in terms of the momenta. The wave equation is to be derived by setting up that differential equation of the second order, linear in the second order derivatives, which has the moving surfaces $F(q, t) = 0$ as its wave-fronts, or, as the mathematician says, its characteristics.[1] The point we wish to make in this note is that while the first order, second degree equation giving the characteristics of a second order, first degree equation are definitely determined by this latter equation the converse is by no means true; there are many second order, first degree equations possessing the same characteristics; or, put in a form more familiar to the physicist, there are many wave equations possessing the same wave-fronts. Merely for the sake of brevity and convenience in writing we shall confine our discussion to the case of a dynamical system with one positional coördinate x and shall indicate how the wave-fronts of an equation of the form

$$g \frac{\partial^2 \psi}{\partial x^2} - \frac{\partial^2 \psi}{\partial t^2} + \varphi = 0$$

are determined. Here g and φ may be any functions of the independent variables (x, t), the dependent variable ψ, and its two partial derivatives $p_x = \frac{\partial \psi}{\partial x}$, $p_t = \frac{\partial \psi}{\partial t}$.

The family of moving wave-fronts $F(x, t) = 0$ of this wave-equation may be conveniently represented as a family of curves in an (x, t) plane. Let us suppose the initial wave-front as represented by a curve whose equation is $t = f(x)$. Then if we assume that the disturbance is travelling in the positive direction along the x axis, points x to the right of this curve, having not yet been reached by the disturbance which is described by

the values of ψ, must have ψ equal to zero; hence ψ and all its derivatives with respect to x and t vanish to the right of the curve $t = f(x)$. If the forces causing the disturbance are not impulsive in character we must assume continuity of ψ and its first time derivative so that we have both ψ and $\frac{\partial \psi}{\partial t}$ equal to zero all along the curve $t = f(x)$. We shall assume our wave-equation homogeneous so that it does not have in it any term independent of ψ; for the sake of argument let us take φ linear in ψ and its first derivatives, i.e.,

$$\varphi = ap_x + bp_t + c\psi; \quad (a, b, c) \text{ functions of } (x, t).$$

Then an obvious solution of the problem is the trivial one $\psi = 0$ for all values of (x, t); the hypothesis that an actual disturbance is being propagated tells us that the problem as set up is not definite, there being at least two answers one of which we have rejected as trivial. This indeterminateness in the problem gives us information as to the shape of the curve $t = f(x)$ in the following way. We are given that ψ and p_t are each equal to zero, identically in x, along the curve $t = f(x)$. Differentiating the identity $\psi = 0$ with respect to x we obtain $p_x + p_t t' = 0$ where t' denotes the derivative of t with respect to x along the curve $t = f(x)$. Since $p_t = 0$ this tells us that $p_x = 0$ along this curve. Again differentiating $p_t = 0$ and $p_x = 0$ with respect to x we find $p_{tx} + p_{tt}t' = 0$; $p_{xx} + p_{xt}t' = 0$ so that $p_{xx} = p_{tt}(t')^2$ along the curve $t = f(x)$. We have here used the easily understood notation p_{xx}, etc., for $\frac{\partial^2 \psi}{\partial x^2}$, etc. Upon substitution of this value of p_{xx} in our equation $gp_{xx} - p_{tt} + \varphi = 0$ we have for the determination of p_{tt} the equation

$$p_{tt}[g_0 t'^2 - 1] + \varphi_0 = 0.$$

In order that p_{tt} may be indeterminate it is necessary that both $g_0 t'^2 - 1$ and φ_0 should vanish identically in x; g_0 and φ_0 denoting the result of substituting the value zero for ψ, p_x, p_t wherever they occur in g and φ, respectively. In the particular case we are proposing to discuss $\varphi = ap_x + bp_t + c\psi$ so that $\varphi_0 = 0$ and we see that there is indetermination in the result when $g_0 t'^2 - 1 = 0$. If the wave-front is given in the more general form $F(x, t) = 0$ we have $\frac{\partial F}{\partial x} + \frac{\partial F}{\partial t} t' = 0$ so that the wave-fronts must satisfy the equation

$$g_0 \left(\frac{\partial F}{\partial x}\right)^2 - \left(\frac{\partial F}{\partial t}\right)^2 = 0.$$

Conversely if we are given the information that the wave-fronts satisfy this equation all we can say is that the wave-equation is of the form

$$gp_{xx} - p_{tt} + \varphi = 0$$

where g and φ are, in the general case, functions of (x, t, ψ, p_x, p_t) which reduce, respectively, to g_0 and zero when each of the three variables (ψ, p_x, p_t) is set equal to zero. In the cases likely to be of interest g will involve only the independent variables (x, t), so that $g_0 = g$, and φ will be of the form $ap_x + bp_t + c\psi$ where (a, b, c) are functions of (x, t). The introduction of these terms involving the first derivatives and the dependent variable itself has some interest in connection with Schrödinger's generalization of his wave-equation in order to take account of a magnetic field.[2]

In order to get some idea of the physical meaning behind the addition of the terms involving *the unknown function* ψ and its first derivatives with respect to the independent variables (x, t) to the wave equation let us see how arbitrary is the assignment of values to the indeterminate second derivatives over the wave-front. Denoting, as before, by primes differentiation with respect to the single independent variable x on the wave-front we have the equations

$$p_t' = p_{tx} + p_{tt}t'; \quad p_{tx}' = p_{txx} + p_{ttx}t'; \quad p_{xx}' = p_{xxx} + p_{txx}t'$$

which enable us to express p_{xxx} and the other third order derivatives in terms of p_{ttt}. We assume here that ψ is an analytic function of (x, t); physically this implies that the forces are continuous. On differentiating our wave equation with respect to x we find

$$gp_{xxx} - p_{ttx} + g_x p_{xx} + \varphi_x = 0$$

where g_x, for example, means the complete derivative of g with respect to x, t being kept constant; if g involves ψ, for instance, the rate of change of ψ with respect to x will have to be taken care of in setting up g_x. Replacing p_{xxx} and p_{ttx} by their expressions in terms of p_{ttt} we have

$$g_x p_{xx} + gp_{xx}' - gp_{tx}'t' + gp_{tt}'t'^2 - [gt'^2 - 1]p_{ttt}t' - p_{tt}' + \varphi_x = 0.$$

The coefficient of p_{ttt} is zero on the wave-front; further, since the values assigned to p_{tt} on the wave-front must be such as to make $gp_{xx} - p_{tt} + \varphi$ identically equal to zero over the wave-front, since this expression is the left-hand side of our fundamental differential equation, we have on differentiation with respect to x (now the only independent variable) the relation

$$(g_x + g_t t')p_{xx} + gp_{xx}' - p_{tt}' + \varphi_x + \varphi_t t' = 0.$$

Combining these two results we have, since t' is not identically zero,

$$g_t p_{xx} + gp_{tx}' + \varphi_t - gp_{tt}'t' = 0.$$

Using the relations already found, namely, $p_{xx} = p_{tt}t'^2$; $p_{tx} = -p_{tt}t'$, $gt'^2 = 1$ this reduces to

$$(\log p_{tt}^2 t')' = \left(\frac{\partial}{\partial t}\log g\right)t' + \left(\frac{\partial \varphi}{\partial t}\right)t'/p_{tt}$$

writing $\varphi = ap_x + bp_t + c\psi$ we have $\dfrac{\partial \varphi}{\partial t} = a_t p_x + ap_{tx} + b_t p_t + bp_{tt} + c_t\psi +$ cp_t and this, on the moving wave-front where ψ and its first derivatives vanish, has the value $ap_{tx} + bp_{tt} = (b - at')p_{tt}$. Hence

$$(\log p_{tt}^2 t')' = \left(\frac{\partial}{\partial t}\log g + b - at'\right)t'.$$

This means that p_{tt} is completely determined save for a constant factor. The development of ψ in a power series about a point (x, t) on the wave-front then tells us that ψ is completely determined save for a constant factor. In the applications already made of Schrödinger's theory g has been a function of x alone so that our equation for the distribution of p_{tt} over the wave-front is

$$(\log p_{tt}^2 t')' = (b - at')t'.$$

If the terms in the first derivatives are not present in the wave-equation, i.e., if a and b are zero this gives $p_{tt}^2 t'$ constant so that p_{tt} (or in other words the acceleration on the wave-front) is proportional to the square root of the velocity of propagation of the wave-front. It may be observed that if $b = at'$, i.e., if $a = b\sqrt{g}$ the values of p_{tt} on the wave-front are not altered by the presence in the wave equation of the terms involving the first derivatives.

B. *The Energy-Frequency Relation.*—In Schrödinger's paper[3] the equation

$$E = h\nu$$

is either introduced as a separate explicit assumption or by assuming that the wave-function ψ depends on W only. There seems to be a way to deduce the relation and the following indicates the line of ideas.

First Schrödinger shows that there will be a geometrical correspondence between the paths of particles and the "rays" or orthogonal trajectories of wave surfaces, if the wave-equation is suitably formed; it is

$$\Delta\psi = \frac{2m(E-V)}{E^2}\frac{\partial^2\psi}{\partial t^2}.$$

While the geometrical behavior is the same, the phase velocity of the wave is given by $u = \dfrac{E}{\sqrt{2m(E-V)}}$ but the particle moves with the velocity

$v = \sqrt{2(E-V)/m}$. This is just as it was in the relation between Newton's corpuscular theory of light and Huyghens' wave theory, as was pointed out by Hamilton. The difference between u and v is really the basis of the experimental discrimination between the corpuscular and the wave-theory of light. If we have plane waves falling on the plane surface of a denser medium the geometrical behavior (refraction toward the normal) can be described by assuming a sudden decrease in the value of V at the surface of separation between the two media. Hence u should be smaller while v should be greater in the denser medium. Fizeau's experiments on the velocity of light gave a lower value in water, thus deciding for u instead of v as the velocity of light. What Fizeau really measured was the group velocity of the light waves but the known dispersion in the case of light waves makes the group velocity u' not very different from the phase velocity u. In the new wave-theory of mechanics, on the other hand, such a strong dispersion is assumed that the group-velocity u' behaves like v instead of behaving like u.

In the optical case we have really only the wave-fronts, which, it is true, can be cut into parts but in the case of the new quantum mechanics there is actually something present which acts as if a small particle moved with the velocity v, if we restrict ourselves to the consideration of "macroscopic" motions. We have, therefore, to build up a wave-motion behaving like a particle (what Schrödinger calls a wave-packet). This can be done, as Schrödinger has pointed out,[4] by superimposing monochromatic plane waves. The wave-packet will move with the group-velocity u' which is given by the classical formula $u' = \dfrac{d\nu}{d(\nu/u)}$. In our wave equation the energy constant E is the only thing that can be connected with the frequency ν (the potential energy V being regarded as determined by the space coördinates). Let us regard E as a function of ν and ask ourselves what functional relationship will make the group velocity u' of the wave the same as the velocity v of the wave packet. Using the value of u given above we find

$$1/u' = \frac{d}{d\nu}\left(\frac{\nu}{u}\right) = \frac{\sqrt{2m(E-V)}}{E} + \nu\frac{dE}{d\nu}\left(\frac{1}{2E}\sqrt{\frac{2m}{(E-V)}} - \frac{\sqrt{2m(E-V)}}{E^2}\right).$$

On putting in for u' the value given for v this leads, after a little reduction, to $dE/E = d\nu/\nu$ whence

$$E = h\nu$$

where h is a constant of integration. de Broglie[5] showed that this relationship is sufficient to make the group velocity of the wave the same as the velocity of the packet; the argument given here shows that the relation is not merely sufficient but necessary.

In the case of restricted relativity theory we have

$$u = \frac{cE}{\sqrt{(E-V)^2 - m^2c^4}}; \quad v = c\sqrt{1 - \frac{m^2c^4}{(E-V)^2}}$$

where c is the velocity of light. The same reasoning as above leads to the same result. Epstein[6] has just published a paper which leads to the relation $E = h\nu$ along somewhat different lines.

[1] Goursat, E., *Cours d'Analyse*, Tome, **3**, Chap. 25.
[2] *Ann. Physik*, **81**, p. 133, 1926.
[3] Schrödinger, E., *Ibid.*, **79**, p. 489, 1926.
[4] Loc. cit., p. 500ff.
[5] de Broglie, L., Theses, Paris, 1924; *Ann. physique* (10), **3**, p. 38, 1925.
[6] Epstein, P., these PROCEEDINGS, **13**, 94, 1927.

HEATS OF CONDENSATION OF POSITIVE IONS AND THE MECHANISM OF THE MERCURY ARC

BY K. T. COMPTON AND C. C. VAN VOORHIS

DEPARTMENT OF PHYSICS, PRINCETON UNIVERSITY

Communicated April 12, 1927

The two most suggestive lines of approach to the problem of accounting for the current at the cathode of a mercury arc are based on considerations of space charge and of thermal equilibrium.[1] In this paper we wish to point out the significance of some recent work by Güntherschultze[2] on evaporation and conduction heat losses from a cathode and by ourselves[3] on heats of condensation of electrons and positive ions. These latter are important factors in the "energy balance" at the cathode, since the cathode is cooled by the emission of electrons from it and heated by the neutralization of positive ions at its surface.

The heat of evaporation of electrons φ_e was first measured by Wehnelt and Jenstzsch[4] and their heat of condensation by Richardson and Cooke.[5] That of positive ions has never been measured, but has been estimated from φ_e, the ionizing potential V_i and the latent heat of condensation of the neutral gas L to be[6]

$$\varphi_+ = V_i + L - \varphi_e. \qquad (1)$$

This equation is obtained from the following simple cycle (Fig. 1). A positive ion at I may condense directly liberating heat φ_+. Or it may do so as at II in the following series of operations: (1) an electron evapo-

rates from the metal, absorbing heat φ_e; (2) it and the ion recombine, liberating heat V_i; (3) the resulting neutral atom condenses in the surface liberating heat L. Comparison of the two equivalent processes leads directly to (1).

While this must be correct as regards the total heat evolved, it does not follow that all this heat must appear as heating of the electrode, since some of it may be radiated away. For instance, if the process of neutralization of the ion occurs at or just outside the surface of the electrode, half the energy is at once lost by radiation and some of the remainder is lost by reflection. Hence the actual heating of the metal φ_+ is less than given by equation (1) and may conceivably be negative. To test this matter, φ_+ was experimentally measured as follows:

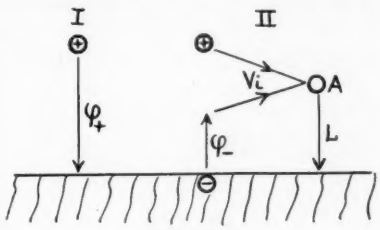


FIGURE 1

A low voltage arc in a gas (A, N_2, H_2 or Hg) was maintained between a tungsten filament cathode C and an anode A, thus creating an atmosphere of intense ionization. In this ionized gas was placed a small sphere S of the metal under investigation (Mo, Pt, etc.), which was supported by three fine leads, of which two constituted a thermojunction T to measure the temperature of the sphere and the third L served to carry the current of electrons and ions to it.

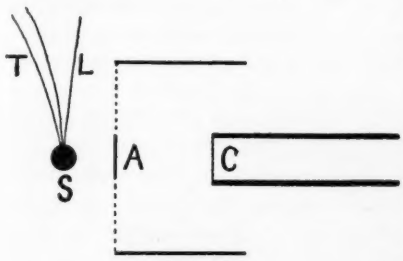


FIGURE 2

Three principal steps were involved in the experiment. First the sphere was used as a Langmuir exploring electrode[7] to determine the mean kinetic energy of agitation of the electrons and the potential of the space in its immediate neighborhood. In principle this consisted in plotting the logarithm of the electron current to the sphere as a function of its potential, which gave a curve like figure 3 in which the discontinuity D indicates the space potential V_s and the slope gives the mean energy $\overline{V}_e$ with which the electrons reach the sphere when V is less than V_s. Second, the current was set at some value, as at A, and the temperature allowed to come to equilibrium, after which the current was suddenly increased by Δi to the value B and the resulting rate of temperature rise measured. The increase of energy input which resulted from this change was

$$\frac{dH}{dt} = ms\frac{dT}{dt} = \Delta i(\bar{V}_e + \varphi_e), \tag{2}$$

where the heat capacity ms was known from the mass m and specific heat s of the sphere, the rate of temperature increase dT/dt due to current increment Δi was measured and $\bar{V}_e$ was known from the preceding measurement. Thus the heat of condensation of electrons φ_e in the metal was at once given. Third, the potential of the sphere was set at a more negative value such that the numbers of electrons and positive ions reaching it were exactly equal and the net current was zero. The temperature attained by the sphere in this condition was measured. Then that larger negative potential was found at which the sphere attained the same temperature, i.e., at which the reduction in heating due to the reduced number of electrons reaching the sphere was just compensated by the increased heating due to the increased number of positive ions and the increased field through which they fell onto the sphere. From these two voltages and the previously determined values of $\bar{V}_e$ and φ_e, together with certain minor corrections which need not be discussed here, the corresponding quantity $(\bar{V}_+ + \varphi_+)$ for positive ions was calculated. The full details of these measurements are being published elsewhere.[3] The values of φ_e thus given are extremely accurately reproducible, and yield interesting information regarding the influence of various gases and methods of surface treatment. The value of $(\bar{V}_+ + \varphi_+)$ is much less accurate, for it emerges from the calculations as the difference between two much larger quantities. Yet the value is quite definite enough to justify the following considerations.

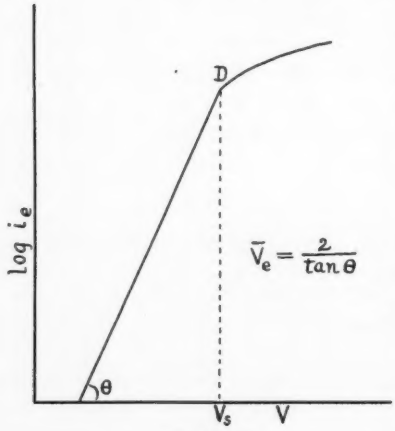


FIGURE 3

Equation (1) applied to argon ions neutralized at the surface of molybdenum leads to $\varphi_+ = 15.3 + 0 - 4.7 = 10.6$ volts, expressing all the quantities in equivalent volts. The heating effect of a positive ion is, therefore, given as more than twice that of a condensing electron. Our experiments, however, consistently gave $(\bar{V}_+ + \varphi_+)$ between 1 and 2 volts. $\bar{V}_+$ is known to be considerably smaller than $\bar{V}_e$ and hence probably less than 0.5 volt, whence φ_+ must be of the order of 1 volt. This decidedly supports

the suggestion that a large portion of the energy V_i is radiated away in the process of neutralization of the ion at the surface. We should expect 50% to be radiated directly away from the metal, and perhaps 25% of the remainder to be reflected from its surface, leaving only about 5.7 out of 15.3 to represent energy of V_i actually supplied to the metal. Thus we should expect about $\varphi_+ = 5.7 + 0 - 4.7 = 1.0$, which is as near the experimental value as could be expected in view of our ignorance of the reflecting power of the metal under the conditions here involved, and the fact that the experimental value is uncertain within a probable range of a volt or so.

It, therefore, seems established for argon and molybdenum that equation (1) for the heating effect of a positive ion should be changed to

$$\varphi_+ = rV_i + L - \varphi_e \qquad (3)$$

in which r is less than 0.5 by an amount depending on the reflecting power of the electrode. There is every reason for extending this to other gas-metal combinations, though we are not yet ready to report on further experimental values. It is evident that φ_+ is, in general, very small compared with the value predicted by equation (1) and that it may even assume negative values.

We now turn our attention to the result of using φ_+ as given by equation (3) in place of equation (1) in calculating the energy balance at the cathode of an arc. In the cases of arcs whose electron emission from the cathode is primarily of thermionic origin (such as the carbon and tungsten arcs considered by Compton[1]) this correction makes little difference in the results, since the condensation of positive ions already contributed but a small part of the total heating. In the case of the mercury arc, however, the case is quite different. We shall base our discussion on data given by Güntherschultze[8] and corrected by Seeliger,[9] which are more accurate than the data used earlier by Compton.[1]

A. Energy lost by cathode:

		Watts per Amp. of Arc Current
(1)	By conduction from cathode spot and ultimately lost to surroundings	2.68
(2)	By evaporation of mercury from cathode (uncertainty due to lack of knowledge of temperature of escaping vapor)	2.8 to 3.9
(3)	By radiation from cathode spot	0.04
(4)	By emission of electrons, if f is the fraction of total current at the cathode which is carried by electrons	$f\varphi_e$
	TOTAL LOSS	$5.52 + f\varphi_e$ to $6.62 + f\varphi_e$

B. Energy gained by cathode:

		Watts per Amp. of Arc Current
(5)	By impact of positive ions which have fallen through cathode drop	$(1-f)V_c$
(6)	By condensation of positive ions	$(1-f)\varphi_+$
(7)	Returned from electrons	$[fV_c-(1-f)V_i]F$
	TOTAL GAIN	$(1-f)(V_c+\varphi_+)+[fV_c-(1-f)V_i]F$

In (7) fV_c is the total energy acquired by the electrons in the fall space, $(1-f)V_i$ is the energy of these electrons which is utilized in ionization, and F is the fraction of the remainder which returns to the cathode in the form of radiation, etc.

Equating A and B gives for equilibrium

$$(5.5 \text{ to } 6.6) + f\varphi_e = (1-f)(V_c+\varphi_+) + [fV_c - (1-f)V_i]F. \quad (4)$$

Substituting $\varphi_e = 3.9$ (Güntherschultze[2]), $V_i = 10.4$, $\varphi_+ = 0$ (from equation (4) with r assumed to have the same value as for molybdenum) and $V_c = 8.6$, and assuming $F = 0$, we find the fraction of current carried by electrons to be $f = 0.25$ to 0.16. If F is taken to be greater than zero, f becomes still smaller.

But it is obvious that if $V_c = 10.4$, f must exceed 0.5, for it could be as small as 0.5 only if the probability of ionization at 10.4 volts were unity so that every electron produced a positive ion. Actually V_c is almost certainly less than 10.4 and the probability is less than unity, both of which necessitate a value of $f > 0.5$ (and probably considerably greater) in order to account for the maintenance of the current. Hence our value of f calculated above is certainly far too small, which forces us to the following reconsideration of the assumptions underlying equation (4).

(*a*) The value of V_c is not known accurately. Experiments by Stark[10] gave 5.3 by the old sounding electrode method, which in the positive column gives a value about 5 volts too small.[7] Just beyond the cathode fall space the error is not so large (McCurdy[11]) so that V_c lies certainly between 5.3 and 10.3. No value between these limits, when introduced into equation (4), gives a possible value of f.

(*b*) The heat lost by evaporation in $A(2)$ above may be over-estimated, since the mercury may escape from the cathode largely as a spray rather than as evaporating atoms. But even if this heat loss is neglected altogether, equation (4) gives only $f = 0.47$, which is still too small.

(*c*) If the electron emission is, as Langmuir[12] suggests, due to actual or partial pulling out of electrons from the cathode by the intense field due to positive ion space charge, then the cooling effect is no longer φ_e, but less. If the emission is due entirely to this process φ_e vanishes and

equation (4) gives $f = 0.56$, which is a possible value, but still improbably small.

(*d*) If effects (*b*) and (*c*) coexist, f may be as large as 0.78, which is quite a reasonable value.

None of the remaining factors in equation (4) seem capable of much alteration or are able to account for the necessary value of f.

From this study we are, therefore, driven to two very important conclusions in the theory of the mercury arc: (1) Langmuir's theory that electrons are drawn out by the intense space charge field is correct. (2) The mercury lost by the cathode is in part lost in lumps or drops composed of numbers of atoms.

[1] K. T. Compton, *Phys. Rev.*, **21**, 226, 1923.
[2] Güntherschultze, *Zeits. Phys.*, **31**, 509, 1925.
[3] To be published by C. C. Van Voorhis in the *Physical Review*.
[4] Wehnelt and Jenstzsch, *Verh. D. Phys. Ges.*, **10**, 610, 1908.
[5] Richardson and Cooke, *Phil. Mag.*, **20**, 173, 1910; **21**, 404, 1911.
[6] Schottky and von Issendorf, *Zeits. Phys.*, **26**, 85, 1924; K. T. Compton, see Ref. 1.
[7] Langmuir and Mott-Smith, *G. E. Rev.*, **27**, 449, 538, 616, 762, 810, 1924.
[8] Güntherschultze, *Zeits. Phys.* **11**, 74, 1922.
[9] Seeliger, *Phys. Zeits.*, **27**, 22, 1926.
[10] Stark, *Ann. d. Phys.* **12**, 1, 1903.
[11] McCurdy, *Phil. Mag.*, **48**, 898, 1924.
[12] Langmuir, *G. E. Rev.*, **26**, 731, 1923.

SERIES SPECTRA OF SILVER-LIKE ATOMS

BY R. J. LANG

UNIVERSITY OF ALBERTA, EDMONTON

Communicated March 21, 1927

The series spectra of the stripped atoms homologous with silver were investigated by Carroll[1] who identified the first members of the Principal, Diffuse and Fundamental Series of In III and the first Principal pair of Sn IV. It has now been possible to complete the identification of the first members of each of the four ordinary series for In III and Sn IV and to determine these also for Sb V.

The investigation was carried out by means of a vacuum spectrograph fully described elsewhere,[2] employing a grating of two meters radius and 30,000 lines per inch. The high-potential spark in vacuo was used as source. The standards of wave-length employed were those obtained by Mr. Smith and the author.[3] The plates obtained were of excellent definition for indium and tin especially (see Plate I); those of antimony, while not so strong, were of fair definition. Carroll's choice of the first Principal

pairs of In III and Sn IV by means of the Doublet Laws was first confirmed; also his Diffuse lines for In III. The Fundamental pair was also photographed again and the satellite accompanying the larger line was found. The separation confirms the measurements made on the Diffuse

TABLE 1

SERIES LINES OF IN III

SERIES NOTATION	λ(I. A. VAC.)	I	ν	Δν	Δλ
$5^2S_1-5^2P_2$	1748.71	25	57185	4342	123.4
$5^2S_1-5^2P_1$	1625.31	20	61527		
$5^2S_1-6^2P_2$?	691.82	1	144546	1335	
$5^2S_1-6^2P_1$?	685.49	1	145881		
$5^2P_1-6^2S_1$	1530.17	9	65352	4345	
$5^2P_2-6^2S_1$	1434.78	8	69697		
$5^2P_1-7^2S_1$	926.89	2	107888	4343	
$5^2P_2-7^2S_1$	891.02	1	112231		
$5^2P_1-5^2D_3$	1494.08	10	66931	290	
$5^2P_1-5^2D_2$	1487.63	12	67221	4343	
$5^2P_2-5^2D_3$	1403.03	10	71274		
$5^2D_3-5^2F_4$	3009.96	2	33223	8.6	
$5^2D_2-5^2F_3$	3008.31	14	33231	291.3	
$5^2D_3-5^2F_4$	2982.93	12	33514		

group. Carroll's wave-lengths were used for the main pair as standards were not available on the plate. The Sharp pair for In III was then located by means of the separation and in the position to be expected by extrapolation from Ag and Cd. This completes the first four groups listed

TABLE 2

SERIES LINES OF SN IV

SERIES NOTATION	λ(I. A. VAC.)	I	ν	Δν	Δλ
$5^2S_1-5^2P_2$	1437.63	30	69564	6513	123.2
$5^2S_1-5^2P_1$	1314.45	20	76077		
$5^2P_1-6^2S_1$	1019.76	10	98062	6509	
$5^2P_2-6^2S_1$	956.29	5	104571		
$5^2P_1-5^2D_3$	1120.76	5	89225	106	
$5^2P_1-5^2D_2$	1119.43	20	89331	6513	
$5^2P_2-5^2D_3$	1044.52	20	95738		
$5^2D_2-5^2F_4$	2089.04	25	47869		
$5^2D_2-5^2F_3$	2084.86	20	47964	107.5	
$5^2D_3-5^2F_4$	2084.38	1	47976	12	
$5^2D_2-6^2F$?	1132.89	2	88270	110	
$5^2D_3-6^2F$?	1131.47	2	88380		

in table 1. The first Diffuse members for Sn IV were next sought and those listed in table 2 were chosen because they fall in the region to be expected and their separation is correct. We find, however, that the D_2D_3 separation is very much smaller than is to be expected by a comparison with Ag I and Cd II. The position of the satellite is correct, however, and the intensity ratio about what would be expected, while no other lines on these plates offer any alternative choice whatever. The choice seems to be further confirmed by the location of the Fundamental pair and satellite in the expected position but with satellite on the short wave-length side of the main pair. The Sharp pair was next located in the positions indicated in table 2.

Proceeding to Sb V, it was first necessary to identify the Principal pair by means of the Doublet Laws by extrapolation from the preceding elements. These were found at 1226.00 and 1104.32 having $\Delta\gamma = 8987$ and $\Delta\lambda = 122$ as listed in table 3. The curve given by the wave number against atomic number for the first Principal lines is very smooth and of nearly uniform curvature and there seems little doubt of their validity. The intensity ratio of the Sb V pair is of the value to be expected. However, it must be stated that intensity ratios are very unreliable for any of these lines which are separated so widely on the plates. There are variations in reflecting power of the grating and sensitivity of the photographic plates involved in all of them. Having located the first Principal pair a search was made for the other groups which resulted in the lines listed in table 3. With the exception of the Sharp pair in which the intensity ratio is definitely wrong, the identification of these lines was straightforward. The plates of antimony, as stated above, were not so sharp and strong as for the other elements and the accuracy of measurement is not so high, so that in these short wave-lengths the variations in $\Delta\gamma$ are probable within experimental error.

TABLE 3

SERIES LINES OF SB V

SERIES NOTATION	λ(I. A. VAC.)	I	γ	Δγ	Δλ
$5^2S_1-5^2P_2$	1226.00	12	81566	8987	121.7
$5^2S_1-5^2P_1$	1104.32	8	90553		
$5^2P_1-6^2S_1$	746.06	1	134037	8980	
$5^2P_2-6^2S_1$	699.22	3	143017		
$5^2P_1-5^2D_3$	898.02	1	111356	826	
$5^2P_1-5^2D_2$	891.41	6	112182	8981	
$5^2P_2-5^2D_3$	831.00	6	120337		
$5^2D_2-5^2F$	1524.47	12	65595	820	
$5^2D_3-5^2F$	1505.70	12	66415		

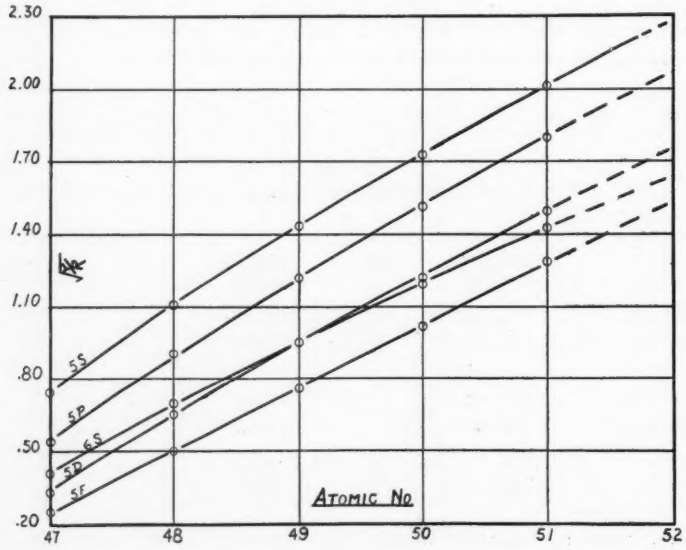
The term values listed in table 4 were arrived at by treating the f orbits as hydrogenic in the usual way and obtaining thus an estimate of the F

TABLE 4

TERM VALUES

	5^2S_1	5^2P_1, 5^2P_2	5^2D_2, 5^2D_3	$5F_{3,4}$	6^2S_1	7^2S_1
Ag I	61096	30623	12331	6892	18540	
		31543	12352			
Cd II	136377	89758	46531	27944	53386	
		92241	46685			
In III	227230	165704	98483	65255	100350	57816(?)
		170047	98773			
Sn IV	327276	251200	161868	114022	153141	
		257712	161974	114010		
Sb V	449300	358748	246566	180970	224711	
		367735	247391			

term value. From this the other term values were calculated. Then a curve was plotted between $\sqrt{\gamma/R}$ for each term, as given in table 5, against Atomic number. These curves are shown in figure 1 from which it will be seen that the curves are smooth and those for $5S$ and $5P$ are parallel;



also that the $5D$ curve and the $6S$ curve cross each other at about the second stage of ionization. The actual values of these terms may still be in error by a considerable amount.

When the work in hand had reached this stage a plate or two of tellurium became available and while these were not of very good quality (see

TABLE 5

$\sqrt{\gamma/R}$ FOR TERMS IN TABLE 4

	5^3S_1	5^3P_1, 5^3P_2	5^3D_2, 5^3D_3	$5^3F_{3,4}$	6^3S_1
Ag I	0.744	0.528	0.335	0.251	0.411
		0.536			
Cd II	1.115	0.904	0.651	0.505	0.698
		0.917			
In III	1.434	1.223	0.941	0.760	0.951
		1.239			
Sn IV	1.727	1.513	1.214	1.019	1.181
		1.532			
Sb V	2.024	1.808	1.499	1.284	1.431
		1.831			

plate) they were sufficient to locate the First Principal pair as shown in table 6. It will be seen that $\Delta\lambda$ has the expected value of 123.

TABLE 6

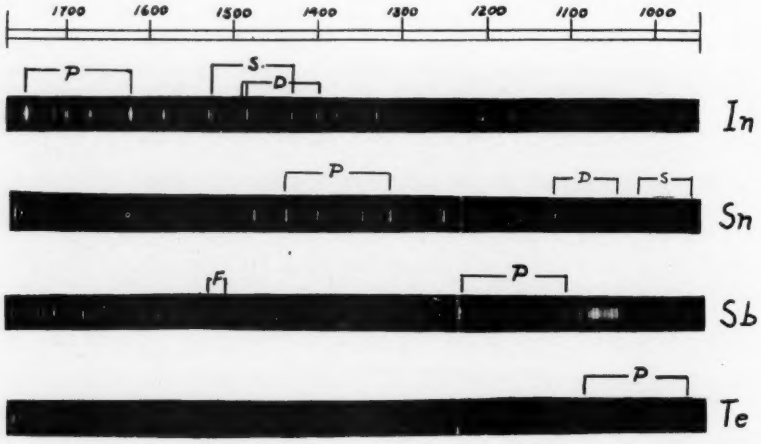
SERIES LINES IN TE VI

SERIES NOTATION	λ(I. A. VAC.)	I	γ	Δγ	Δλ
$5^2S_1-5^2P_2$	1077.91	8	104760	11988	123.3
$5^2S_1-5^2P_1$	954.57	5	92772		

Second Members of Series.—It is known that the high-potential spark source does not generally produce higher members of these series of the elements. However, in some cases these do appear as in the author's work[4] on Ca^+ in which higher members of the F series occur. It was evident that if these should appear on the plates used those for Indium would be most favorably placed for observation by means of the present apparatus and hence a search was made for these which quickly resulted in the location of a pair of lines at 926.89 and 891.02 which appear to fit the requirements for the second pair of the Sharp Series, since the separation is exact to one unit of wave number and the position within 10 Å of that calculated by a Rydberg formula using as limit the rough value as obtained by the method already described. With these two pairs of lines taken as the first two members of the Sharp series a new calculation of $5P$ was made and all other terms were corrected from this resulting in the values given for indium in table 4. This correction was of the nature of a slight increase in the value of all of the terms and amounted to considerably less than one per cent in the $5S$ term.

Having obtained probably better measures of the term values the second members of the other series of indium were calculated. The results

of this were not so satisfactory although the pair listed at 691.82 and 685. may possibly be the second pair of the P series. The second F pair should occur close to the strong line at 1748.71, and while two lines were found here one on each side of the strong line and having $\Delta\gamma = 90$ these have been omitted because one of them appears to be a ghost of the main line.



Second members of the series of Sn IV will, in general, lie rather far into the extreme ultra-violet with the exception of the F series which should occur at about 1100 Å. The pair listed seems to fit the requirements rather well. The separation seems to be slightly larger than the D levels but the difference is just on the limit of experimental error. The intensity ratio seems to be correct. Another grating will be used shortly to examine the region of shorter wave-lengths for higher members of the other series of tin.

The author wishes to express his appreciation to the Honorary Advisory Council for Scientific and Industrial Research of Canada for a grant which enabled him to carry out this research.

[1] *Phil. Trans.*, **A225**, pp. 357–420.
[2] *J. O. S. A.*, **12**, no. 5, May, 1926.
[3] *Phys. Rev.*, **28**, no. 1, July, 1926.
[4] *Astrophys. J.*, **64**, no. 3, Oct., 1926.

TERM-STRUCTURE AND ZEEMAN EFFECT OF THE ARC SPECTRUM OF TIN

(Preliminary Report)

BY J. B. GREEN* AND R. A. LORING

JEFFERSON PHYSICAL LABORATORY, HARVARD UNIVERSITY

Communicated April 14, 1927

While constant frequency differences in the spectrum of tin have been known for some time, the interpretation of the energy levels involved has been an open question. The numerical relationships were quite completely worked out by Sponer[1] from data obtained by McLennan, Young and McLay.[2] More recently, Zumstein[3] has given a very detailed study of the absorption spectrum of tin vapor which has helped materially in the unravelling of the spectrum.

The two outside electrons of tin are in 5_2 orbits and consequently, according to Hund,[4] the lowest energy-levels should be $^3P_{0,1,2}$, $^1\overline{D}_2$, and $^1\overline{S}_0$, the $^1P_1$, $^3\overline{D}_{1,2,3}$, and the $^3\overline{S}_1$ being excluded by the Pauli-Heisenberg exclusion principle. If one electron remains in a 5_2 orbit while the other moves to a 6_1 orbit, we get a $^3\overline{P}_{0,1,2}$ and a $^1\overline{P}_1$. If one electron stays in a 5_2 orbit and the other moves to a 5_3 orbit, we get $^3\overline{F}_{2,3,4}$, $^3D_{1,2,3}$, $^3\overline{P}_{0,1,2}$, and $^1\overline{F}_3$, $^1D_2$, and $^1\overline{P}_1$. This has been noted by Sur,[5] and table 1 is but a slight modification of his table.

The very strong absorption line 3034.16 has not been incorporated in any of the schemes. Zumstein[3] thinks it is a combination between a ^5D term and a ^5P term. It has been placed in the scheme as a $^3P_1{}^3\overline{P}_0$ combination, which is justified by its Zeeman pattern, although the separation $^3\overline{P}_0 - {}^3\overline{P}_1$ is very small. The same assignment has been given to the line 2091.58, also a strong absorption line, assigned by Zumstein[3] to the $^3P_2$ level, but somewhat in doubt.

It is to be noted that most of the lines formed by combinations with $^3P_{0,1,2}$, are strong or moderate absorption lines with the exception of λλ 2790 and 2433 which show no absorption. All the lines formed by combination with $^1\overline{D}_2$ are faint absorption lines and all the lines formed by combination with $^1\overline{S}_0$ were not obtained as absorption lines. This supports the assumption that ^3P is the lowest set of energy levels.

Van der Harst[6] has studied the Zeeman effect of tin and several other metals, but his results are very incomplete. He records merely triplets and quartets. For example, λ 3801 which he calls a triplet is clearly resolvable on the authors' plates in the second order as having nine components. The Zeeman effect was determined for several of the tin lines and the average g values calculated are given in table 1. They are probably correct to about 2%. It is seen that they differ from the theoretical

values, but this difference was also noted by Back[7] in the case of lead, which shows the same fundamental levels as tin, according to Hund's theory.

TABLE 1

	$^3P_0$	%	$^3P_1$	1.43	$^3P_2$	1.46	$^1\bar{D}_2$	1.02	$^1\bar{S}_0$	%	AVER. G.
		Int. =	(6R)S	1.42							
$^3\bar{P}_0$		λ =	*3034.16*	%							%
		ν =	32948.96								
	(6R)S		(6R)S		(5R)S		(4R)f		(5)	%	
$^3\bar{P}_1$	*2863.320*	%	*3009.138*	1.45	*3175.039*	1.46	*3801.031*	1.04	*5631.69*		
	34914.29	1.38	33222.46	1.35	31486.61	1.38	26301.21	1.38	17751.7	1.38	1.375
			(7R)S		(8R)S		(6)f				
$^3\bar{P}_2$			*2706.504*	1.43	*2839.985*	1.46	*3330.596*	1.00			1.43
			36937.10	1.43	35201.13	1.43	30016.04	1.43			
	(5R)S		(4R)S		(2)		(5R)f		(5)	%	
$^1\bar{P}_1$	*2546.552*	%	*2661.245*	1.42	*2790.187*		*3262.338*	1.025	*4524.740*		1.03
	39257.00		37565.23	1.04	35829.35		30644.06	1.025	22094.54	1.05	
			(3R)m		(5R)S		(4R)f				
$^3\bar{F}_2$			2380.742		2483.389		*2850.618*	1.00			0.67
			41990.90		40255.42		35069.84	0.67			
					(6R)S		(4R)f				
$^3\bar{F}_3$					2429.490		*2779.814*	1.02			1.12
					41148.41		35963.03	1.12			
$^3\bar{F}_4$	All combinations excluded by selection principle										
	(6R)S		(5R)S		(2)		(3R)f		(3)	%	
$^3D_1$	2246.02		2334.799		2433.473		2785.027		*3655.78*		
	44509.7		42817.09		41081.01		35896.15		27345.5	0.63	0.63
			(6R)S		(4R)m		(4R)f				
$^3D_2$			2199.29		2286.65		2594.431				
			45454.9		43718.85		38532.59				
					(6R)S		(5R)f				
$^3D_3$					2268.902		2571.598				
					44060.53		38874.70				
			(3R)S								
$^3\bar{P}_0$			2091.58								
			47795.4								
	(2R)		(2R)m		(2R)f		(4R)f		(3)		
$^3\bar{P}_1$	1994.30		2063.95		2140.65		2408.143		3032.783		
	50126.6		48435.1		46700.0		41513.13		32963.45		
			(2R)f		(4R)m		(1)				
$^3\bar{P}_2$			2027.15		2100.83		2357.88				
			49314.3		47585.1		42958.0				
					(6R)m		(5R)f				
$^1\bar{F}_3$					2209.60		2495.722				
					45242.9		40056.50				
			(6R)S		(3)f		(3R)f				
$^1D_2$			2354.840		2455.250		2813.582				
			42452.73		40716.73		35531.08				
	(6R)S		(3R)f		(4R)m		(3R)f		(3)		
$^1\bar{P}_1$	2072.86		2148.44		2231.68		2523.912		*3218.690*	%	1.23
	48226.9		46530.6		44795.3		39609.13		31059.60	1.23	

The medium size of Weiss magnet manufactured by the Société Gènévoise was fitted with a vacuum box placed between the poles of the magnet, and hydrogen was pumped through continuously, maintaining pressure of about 17 mm. The magnet was run at about 80 amperes and for one set of plates measured, the field strength was about 33,500 gausses. The tin was used as the anode and was placed in the rim of a

large brass disk. A brass wire resting on the wheel by gravity served as the cathode. The arc was run at a current of about 4 amperes. This method, devised by one of the authors (R. A. L.) serves to increase very greatly the heat capacity of the anode and the arc consequently runs

TABLE 2

NO.	LINE	OBSERVED PATTERN							CALC. G.	THEOR. G.
1	5632	(0)	1.38	Observed					1.38	3/2
		(0)	1.38	Calculated from g values					%	%
		(0)	1.38	Calculated from aver. g values						
2	4525	(0)	1.05						%	%
		(0)	1.05						1.05	1
		(0)	1.03							
3	3801	(0)	(0.34)	0.70	1.04	1.36			1.04	1
		(0)	(0.34)	0.70	1.04	1.38			1.38	3/2
		(0)	(0.36)	0.68	1.02	1.38				
4	3656	(0)	0.63						%	%
		(0)	0.63						0.63	1/2
		(0)	0.63							
5	3331		(0.40)	(0.86)	0.54	0.97	1.52	1.92	1.00	1
			(0.43)	(0.86)	0.57	1.00	1.43	1.86	1.43	3/2
			(0.41)	(0.82)	0.61	1.02	1.43	1.84		
6	3263	(0)	1.025						1.025	1
		(0)	1.025						1.025	1
		(0)	1.025							
7	3219	Very unsymmetrical				1.19	(0)	1.27	%	%
									1.23	1
8	3175	(0)		1.52					1.38	3/2
		(0)	(0.08)	1.38	1.46	1.54			1.46	3/2
		(0)	(0.08)	1.38	1.46	1.54				
9	3034	(0)	1.42						%	%
		(0)	1.42						1.42	3/2
		(0)	1.43							
10	3009		(0.09)	1.41					1.35	3/2
			(0.10)	1.35	1.45				1.45	3/2
			(0.05)	1.38	1.43					
11	2863	(0)	1.38						%	%
		(0)	1.38						1.38	3/2
		(0)	1.38							
12	2851	(0.33)	(0.71)	0.33	0.71	1.09	Not completely resolved		1.00	1
		(0.33)	(0.67)	0.33	0.67	1.00	1.33		0.67	2/3
		(0.35)	(0.71)	0.33	0.57	1.02	1.37			
13	2840	(0)	1.46						1.43	3/2
		(0.03)	1.43	1.46					1.46	3/2
		(0.03)	1.43	1.46						
14	2780	(0)	1.295		Very fuzzy		Not well resolved		1.02	1
		(0)	(0.10)	(0.20)	0.92	1.02	1.12	1.22 1.32	1.12	12/11
		(0)	(0.10)	(0.20)	0.92	1.02	1.12	1.22 1.32		
15	2706	(0)	1.43						1.43	3/2
		(0)	1.43						1.43	3/2
		(0)	1.43							
16	2661	(0.38)	1.15	1.50			Very difficult to measure		1.04	1
		(0.38)	1.04	1.42					1.42	3/2
		(0.40)	1.03	1.43						
17	2546	Very unsymmetrical					1.01	(0) 1.24	%	%
									?	1

much cooler. Furthermore, no insulation is needed for the magnet poles. This type of vacuum arc which serves very well for low melting-point metals will be described in detail in a later paper. The light from the box was focussed on the slit of the 21-foot concave grating in the Paschen mounting set up in the basement of this laboratory. Exposures were from 1 hr. to 10 hr. and the lines were always sharp since the temperature of the grating room could be maintained constant to within a very few hundredths of a degree by means of an ether thermostat.

Table 2 shows the Zeeman patterns for all the lines measured, together with the I values calculated from them. Especially interesting is the line 3009 which shows no central component, and, therefore, shows that the g value of $^3\overline{P}_1$ is not the same as $^3P_1$, both of which should be $^3/_2$ according to theory. All the assignments of inner quantum numbers are shown to be correct, if we base them on $^3P_0$ as the normal level. The lines λλ 3330, 3009, 2661 and 2851 all have the undisturbed component missing, showing that the transition involved is one between two levels of the same inner quantum number.

Very little will be said at present about the g values. It will be noted, however, that the g values for the $^3P_1$, $^3P_2$, and $^3\overline{P}_1$ and $^3\overline{P}_2$, namely, 1.43, 1.46 and 1.38, 1.43 are very closely $\frac{34}{24}$, $\frac{35}{24}$ and $\frac{33}{24}$, $\frac{34}{24}$.

Several other lines have been obtained and are now being measured and will be published in a later paper. Especially rich in lines is the spark spectrum, which shows some very interesting Zeeman patterns. Regularities are now being searched for in this spectrum which should show a doublet system similar to that of In.

Note Added May 3rd. Except for the five lowest terms in the spectrum and the next four, $^3P_{012}$ and $^1P_1$, the designations as assigned by Sur, shown in table 1 are not definitive. In fact, it is likely that $^1D_2$ and $^3D_2$ should be interchanged, making $^3D_2$ lower than $^3D_1$, which also happens in the case of lead. Also, the last row on the page may be the second member of a series starting with $^3\overline{P}_1$ at the top of the column. These two sets of lines in conjunction with the next three terms of the series, give as the value of the lowest level of tin 59,690, corresponding to an ionization potential of 7.37 volts and a radiation potential of 4.07 volts. Several lines in the infra-red, red and yellow have also been assigned to levels in tin, making the spectra of tin and lead quite homologous.

* NATIONAL RESEARCH FELLOW.

[1] Sponer, *Zeit. Phys.*, **32**, 19, 1925.
[2] McLennan, Young and McLay, *Trans. Roy. Soc. Can.*, **28**, 57, 1924.
[3] Zumstein, *Phys. Rev.*, **27**, 150, 1926.
[4] Hund, *Zeit. Phys.*, **33**, 345, 1925; **34**, 296, 1925.
[5] Sur, *Zeit. Phys.*, **41**, 791, 1927.
[6] Van der Harst, *Arch. Neerl.*, **A11**, 1ff., 1925.
[7] Back, *Zeit. Phys.*, **37**, 193, 1926.